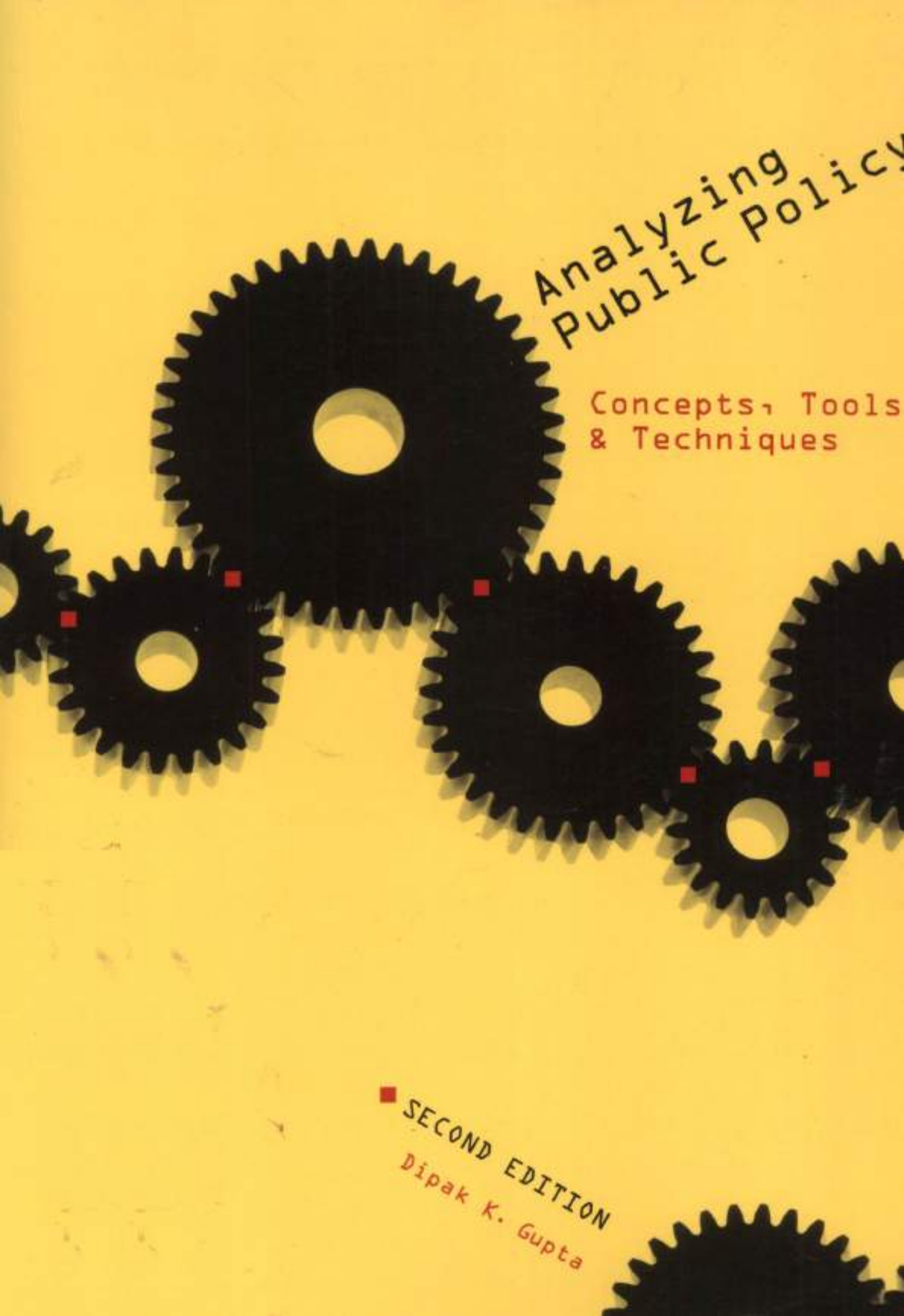




Analyzing Public Policy

Concepts, Tools
& Techniques

■ SECOND EDITION
Dipak K. Gupta

■ INSTRUCTOR'S

RESOURCES, available at <http://college.cqpress.com/instructor-resources/gupta>, include graphics in electronic format and a solutions manual.

■ DIPAK K. GUPTA

is the Fred J. Hansen Professor of Peace Studies, Distinguished Professor in Political Science, and director of a multidisciplinary undergraduate program, International Security and Conflict Resolution (ISCOR), at San Diego State University. The author of eight books and more than one hundred other publications, he specializes in ethnic conflict, terrorism, and public policy.

The author has integrated public policy, policy analysis, and methods into an interesting, well-written, and dynamic book.

JEFFREY GREENE
University of Montana

This is not 'just another methods text.' Gupta presents a coherent, accessible, yet thorough discussion of the major tools of policy analysis. The quantitative focus is tempered with an acknowledgment of the role of politics and other realities in public policy making.

MICHELLE PISKULICH
Oakland University

Gupta has written a comprehensive, competent, and readable text for policy analysis students. Quantitative methods are explained in a clear fashion for those interested in application rather than statistical properties. Examples provide the essential connection between techniques and their applications.

EDWARD MILLER
University of Wisconsin-Stevens Point


CQ PRESS

A Division of SAGE
www.cqpress.com

ISBN 978-1-60426-570-5



9 781604 265705

Analyzing Public Policy

Concepts, Tools, and Techniques

SECOND EDITION

Dipak K. Gupta
San Diego State University




CQ PRESS

A Division of SAGE
Washington, D.C.

CQ Press
2300 N Street, NW, Suite 800
Washington, DC 20037

Phone: 202-729-1900; toll-free, 1-866-4CQ-PRESS (1-866-427-7737)

Web: www.cqpress.com

Copyright © 2011 by CQ Press, a division of SAGE. CQ Press is a registered trademark of Congressional Quarterly Inc.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Cover design: www.designfarm.com
Composition: C&M Digitals (P) Ltd.

© The paper used in this publication exceeds the requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992.

Printed and bound in the United States of America

14 13 12 11 10 1 2 3 4 5

Library of Congress Cataloging-in-Publication Data

Gupta, Dipak K.

Analyzing public policy : concepts, tools, and techniques / Dipak K. Gupta. — 2nd ed.
p. cm.

Includes bibliographical references and index.

ISBN 978-1-60426-570-5 (alk. paper)

1. Policy sciences—Statistical methods. 2. Policy sciences—Mathematical models.
I. Title.

H97.G868 2010
320.6072'7—dc21

2009051866

Analyzing Public Policy

To Chitra and Samar Das
१३

Preface

I began writing the first edition of this book in early 2000. Since then, the world of public policy has felt the shock of several catastrophic national and international events. At the start of the decade, anxiety surrounded the turn of the new millenium (dubbed Y2K), and the stress of a contested presidential election followed shortly. The devastating attacks of 9/11 shocked the world in 2001 and led the United States to protracted wars in Afghanistan and Iraq. Nature demonstrated its fury with a tsunami in Asia in 2004 that killed more than 200,000 people, and with Hurricane Katrina in 2005, whose aftermath exposed grave social problems in the United States. There were concerns about global warming heating up not only the world but also the political climate in America. Finally, in the waning days of the decade, the world experienced the worst economic decline since the Great Depression. No wonder the December 7, 2009 issue of *Time* carried the cover story "The Decade from Hell."¹ These momentous events have led to policy conundrums, exposing the vulnerability of our most strongly held beliefs about what the government should do.

The cover of a July 2009 issue of the *Economist* featured a picture of a melting textbook on "modern economic theory," evoking Salvador Dali's surreal melting clocks, with the caption "Where it went wrong—and how the crisis is changing it."² It is hard to believe that about a year earlier, newly minted economists fresh out of universities were being hired by hedge funds with unlimited money and matching ambition to develop complex models of market behavior based on the traditional assumptions of human behavior. We believed that we had tamed the monster called the "market" by reducing it to a finite number of stylized, interconnected components. Yet, as demonstrated by the economic crisis, the rough edges of reality can cut deeply into our ability to manage the vast outside world.

Meanwhile, however, there was a sea change of ideas taking place. As one part of the academic and policy world remained rooted in tradition, another was changing rapidly. Several disciplines—from behavioral economics and decision sciences to social psychology and cognitive sciences—were shedding new light on the complexities of the human mind. With these advances in knowledge, along with the exposure of the vulnerability of traditional positivism, which sees human beings as "rational" actors, a number of related disciplines have been paving a new way of looking at the policy world. This second edition attempts to sort out this confusion and the complexities of finding the right policy prescription. It emphasizes that policy analysis is less a slavish creed than a prism through which to understand the world. This world is devoid of a single optimal solution and the corresponding ability to engineer all its sides; it is a world in which there is perhaps no right answer, only compromises among conflicting goals, aspirations, interests, and ideologies.

To this end, I have completely revised the first chapter to more fully explore the complex relationship of human behavior to public policy, as well as the social goals that

government must balance in its crafting of policy. Additionally, chapter 2 has undergone a thorough revision; the political and ethical forces in policy analysis receive special consideration as I discuss the tools and structure of the field. I also have added a final chapter that offers some straightforward, practical suggestions for aspiring (and established) policy analysts. Finally, readers will note the inclusion of an excerpt from a Government Accountability Office report on poverty in Appendix A, as well as several end-of-chapter exercises that refer to it. This report serves as a valuable starting point for a real-world policy analysis, and I encourage readers to reflect on the issues I bring up throughout the book as they peruse the report.

In addition to updating and streamlining the content of all chapters, I have added many new “Case in Point” features. In choosing these cases, drawn from all areas of public policy, I have used the dual criteria of relevance and student interest. I have made sure that the cases show the practicality of the topics under discussion, and I have made every attempt to make the narratives of interest to students. Books don’t dream, but their authors do: I fervently hope that these cases generate discussion and debate not only within the structured classroom, but also outside in the hallways, in the parking lots, and in front of the elevators. After teaching for more than three decades, I strongly believe that students excel only when they take an active interest, which is generated when they discover the real-life relevance of what they are learning in classrooms.

Since the aim of this book is to present a comprehensive picture of policy analysis, I had to be mindful of its size and coverage. As a result, I decided to move the old chapter 12, “Models of Causal Prediction: Multiple Regression,” to Appendix B. Multiple regression and causal prediction are introduced in chapter 12 of this edition, but instructors who would like to draw upon the more extensive discussion of these topics from the previous edition will find it in the appendix.

While I have made some serious changes in preparing this second edition, I have not altered the fundamental orientation of the original volume. In the vast marketplace of public policy texts, nearly all of which are written from a particular perspective, this book stands virtually alone in its pursuit of a comprehensive outlook. By looking at numerous syllabi collected from many institutions of higher learning, I have made the prosaic discovery that almost nobody uses a single textbook when teaching public policy; some emphasize the ethical considerations and some the economic principles, some concentrate on the policy process, some stress the methodology of analysis, some focus on statistical tools, and others teach public policy courses as case studies of actual policies. It is my hope that when instructors use this text, they will be able to expand it to their desired ends by incorporating other great books available in the market.

In the construction of my ideas I have benefited a great deal from many of my friends and colleagues and, thanks to the modern technology of the Internet, students from many parts of the world. Among the former group, my intellectual debt to C. Richard Hofstetter, Madhavi McCall, and Ron King runs deep. Among the latter, I have received many constructive comments from students, including one from New Zealand who complained that in the previous edition I did not draw the map of Papua New Guinea accurately. In this current edition, an accurate map of the island has replaced the “glob.” I also benefited from the candid feedback of several reviewers commissioned by CQ Press: Matthew Cahn, California State University, Northridge; Amy Dreussi, University of Akron; Kurt Gaubatz, Old Dominion University; Mark

Imperial, University of North Carolina Wilmington; Robert Stoker, George Washington University; and Jill Tao, University of Hawaii. I am also thankful to James Kelly for an extremely professional editing job.

Finally, I must acknowledge my good luck. A decade ago, my good karma brought an energetic and highly competent editor (any author's dream), Charisse Kiino, to my office for an unscheduled appointment, which resulted in the writing of the book. For this revised edition, my good karma still held; I was ably aided by Allison McKay. Without her active help and support, this version of the book might not have been completed.

Notes

1. Andy Serwer, "The Decade from Hell—and Why the Next One Will Be Better," *Time*, December 7, 2009.
2. "What Went Wrong with Economics," *The Economist*, July 18–24, 2009, www.economist.com/printedition/displayStory.cfm?Story_ID=14031376.

Contents

Tables, Figures, and "A Case in Point"	xv
Preface	xxi

CHAPTER 1 REASON, RATIONALITY, AND PUBLIC POLICY: THE PUZZLE OF HUMAN BEHAVIOR	1
---	----------

Government in Our Lives	1
Politics and Rationality	2
The World of the "Rational Fool"	5
<i>The Illogic of a "Rational Being"</i>	6
<i>The Un-rational Being and Political Behavior</i>	9
Government's Goals and Their Impact on Public Policy	15
<i>Efficiency</i>	16
<i>Equity</i>	18
<i>The Equity-Efficiency Trade-off</i>	19
<i>Liberty</i>	20
<i>Security</i>	21
<i>The Liberty-Security Trade-off</i>	22
Key Words	23
Exercises	24
Notes	24

CHAPTER 2 THE ANALYSTS: THEIR ROLE AND THEIR TOOLS	27
---	-----------

The Prince and the Pundits: Analysts, Academics, and Advocates	28
<i>The Politicos and In-house Policy Analysts: An Uneasy Cohabitation</i>	30
<i>Exit, Voice, and Loyalty: What's an Analyst to Do?</i>	32
The Use of Quantitative Techniques in Public Policy Analysis	35
Structure above a Swamp: Public Policy Analysis and Professional Ethics	40
Key Words	43
Exercises	43
Notes	44

CHAPTER 3 GOVERNMENT AND THE MARKET	47
The Fundamental Contributions of Economic Analyses	49
<i>Marginal Analysis</i>	49
<i>Opportunity Costs</i>	52
Market Failure: Why Government Interferes in a Free Market	52
<i>Lack of Competition</i>	53
<i>Barriers to Entry and Exit</i>	55
<i>Restricted or Asymmetrical Flow of Information</i>	57
<i>Externalities and Social Costs</i>	59
<i>Rising Service Costs: The Appearance of Market Failure</i>	61
Government Failure	62
<i>Inability to Define Social Welfare</i>	63
<i>The Limits of Democracy and the Paradox of Voting</i>	63
<i>Inability to Define the Marginal Benefits and Costs of Public Goods</i>	64
<i>Political Constraints</i>	65
<i>Cultural Constraints</i>	65
<i>Institutional Constraints</i>	65
<i>Legal Constraints</i>	66
<i>Knowledge Constraints</i>	66
<i>Analytical Constraints</i>	66
<i>The Timing of Policies</i>	67
Limiting Government Intervention: Joint Partnership in a Mixed Economy	67
<i>Stimulating the Market</i>	68
<i>Simulating the Market</i>	69
Key Words	72
Exercises	72
Notes	73
 CHAPTER 4 THE POLICY PROCESS	 75
Agenda Setting	76
<i>The Pluralist Model</i>	77
<i>The Elitist Model</i>	81
<i>Impact of Sensational Events</i>	83
<i>Evaluating Costs and Benefits</i>	86
Policy Formulation	87
<i>Explaining Behavior</i>	88
<i>Forecasting Effects</i>	88
<i>Conflict, Inaction, and Nondecision</i>	89
Policy Adoption	91
Policy Implementation	92
Policy Evaluation	93

Policy Change	94
Policy Termination	96
Key Words	97
Exercises	97
Notes	98

CHAPTER 5 CRITICAL THINKING AND RESEARCH DESIGN 101

Five Steps of Objective Analysis	102
Setting Goals	104
<i>Utilitarians and the Pareto Principle</i>	106
<i>The Kaldor-Hicks Compensation Principle</i>	107
<i>The Rawlsian Challenge</i>	108
<i>Setting Public Policy Goals</i>	109
Choosing a Method of Analysis	109
<i>Experimental Design</i>	110
<i>Quasi-experimental Design</i>	112
Choosing the Right Model	115
Forecasting Outcomes and Unintended Consequences	116
Designing Policy Research	117
<i>Limiting the Number of Alternatives</i>	117
<i>Accepting or Rejecting a Project</i>	119
<i>Ensuring the Process Is Fair</i>	121
Challenges to Critical Thinking: Biases in Reasoning	122
<i>Representativeness</i>	122
<i>Availability of Information</i>	125
<i>Adjustment and Anchoring</i>	127
A Few Parting Suggestions	127
Key Words	128
Exercises	128
Notes	128

CHAPTER 6 BASIC STATISTICS 131

Numbers as Storytellers	131
Methods of Descriptive Statistics	131
<i>Measures of Central Tendency</i>	132
<i>Measures of Dispersion</i>	137
<i>Which Measure of Central Tendency to Use</i>	144
<i>A Quick Glance at the Distribution: The Stem-Leaf Method</i>	144
<i>Correlation Coefficient: Pearson's r</i>	145
Key Words	147

Exercises	148
Notes	148
CHAPTER 7 PROBABILITY AND HYPOTHESIS TESTING	149
Objective Probability	149
<i>Probability Distribution</i>	151
<i>Hypothesis Testing and Confidence Intervals</i>	157
<i>The Chi-Square Test</i>	165
Subjective Probability	169
Key Words	171
Exercises	171
Notes	176
CHAPTER 8 SOURCES OF DATA	177
What Are We Measuring?	177
<i>Types of Measurement</i>	177
<i>Accuracy of the Measuring Scales</i>	178
Primary Data: Conducting a Survey	179
<i>Designing a Survey</i>	180
<i>Systematic Errors in Sampling</i>	180
<i>Random Sampling Errors</i>	184
<i>Choosing the Sample Population</i>	185
<i>Choosing the Size of the Sample</i>	186
<i>Choosing Effective Survey Instruments</i>	187
<i>Choosing Polling Methods</i>	189
Quantification of Survey Data	189
Reporting Survey Results	190
Conducting Focus Groups	190
Secondary Data	192
<i>Searching for Information</i>	193
<i>Online Searches</i>	193
<i>Searching the Old-fashioned Way</i>	200
<i>When Data Are Not Available</i>	200
Key Words	201
Exercises	201
Notes	202
CHAPTER 9 MAKING SENSE OF NUMBERS	203
A Picture's Worth: Graphical Methods of Analysis	203
<i>Current versus Constant Dollars</i>	204

<i>Percentage Change</i>	208
<i>Creating an Index</i>	209
<i>Choosing the Type of Graph to Use</i>	210
<i>Graphical Methods in Decision Making</i>	211
To Tell the Truth and Nothing but the Truth	213
<i>Interpretation and Deception</i>	214
<i>Tabular Presentation of Data</i>	221
Those Not-so-Innocent Numbers	222
Key Words	223
Exercises	223
Notes	224

CHAPTER 10 PROJECTION TECHNIQUES: WHEN HISTORY IS INADEQUATE 225

Projection versus Causal Prediction	225
Inadequacy of History	227
<i>Single-Factor Projection</i>	227
<i>Fiscal Impact Analysis</i>	229
<i>Problems of Single-Factor Analysis</i>	230
Judgmental Methods of Projection	232
<i>The Delphi Technique</i>	232
<i>The Feasibility Assessment Technique</i>	238
<i>The Expected Utility Model</i>	242
<i>Origin of the Numbers</i>	244
<i>Shortcomings of the Judgmental Methods</i>	245
Key Words	246
Exercises	246
Notes	247

CHAPTER 11 PROJECTION TECHNIQUES: ANALYSIS OF HISTORICAL DATA 249

The Components of a Data Series	249
The Patterns of Time Trends	253
Adjustment Methods	256
<i>Seasonal Adjustment</i>	258
<i>Trend Adjustment</i>	259
Smoothing Out the Fluctuations	260
<i>Projecting the Immediate Past: Naïve Projection</i>	261
<i>Projecting by the Mean</i>	263
<i>Moving Average</i>	263
<i>Choice of Projection Technique</i>	265

The Politics of Forecasting	265
Key Words	266
Exercises	267
Notes	269

CHAPTER 12 PROJECTION TECHNIQUES: THE METHODS OF SIMPLE AND MULTIPLE LEAST SQUARES 271

The Logic of the Least Squares Method	272
Linear Time Trend: Simple Regression Model	275
<i>Accuracy of the Results</i>	278
<i>High R^2</i>	280
<i>Relevance of the Estimated Coefficients</i>	280
<i>The Significance of Individual Coefficients</i>	283
<i>Presentation of Estimation Results</i>	286
<i>The Number of Observations</i>	286
Trend Changes: Building Multiple Regression Models	287
<i>Abrupt Changes in Trend</i>	288
<i>Abrupt Changes in Slope</i>	292
Gradual Changes in Trend: Estimation of Nonlinear Trends	296
<i>Polynomial Forms</i>	296
<i>Higher Order Polynomials</i>	299
<i>Log-transformed Forms</i>	301
<i>Inverse Forms</i>	302
<i>The Problem of Irrelevant Independent Variables: Adjusted R^2</i>	303
<i>The Significance of Coefficients Taken Together</i>	304
<i>Choosing the Correct Functional Form</i>	305
Forecasting and Its Problems	305
<i>Point Forecasts</i>	306
<i>Interval Forecasts</i>	307
Explaining the Present with the Past: Lagged Dependent Variables	311
Key Words	313
Exercises	314
Notes	317

CHAPTER 13 THE ELEMENTS OF STRATEGIC THINKING: DECISION TREE AND GAME THEORY 319

Getting a Grip on Uncertainty	320
Decision Making and Expected Payoff	321
The Decision Tree	322
<i>Structuring a Decision Tree</i>	324
<i>Evaluating Flood Damage Reduction</i>	325
<i>Risk Tolerance and Expected Payoff</i>	330

Two Active Players: Game Theory	332
<i>Game Theory in Local Government Decision Making</i>	337
<i>The Golden Rules of Decision Making under Uncertainty</i>	338
Strategies to Overcome the Prisoner's Dilemma	339
Other Strategies: Trust and Bargaining	341
Key Words	341
Exercises	341
Notes	343

CHAPTER 14 CHOOSING THE BEST ALTERNATIVE: COST-BENEFIT ANALYSIS 345

Social versus Private Cost-Benefit Analysis	346
Defining Goals	347
Identifying Alternatives	347
Listing the Costs and Benefits of the Alternatives	347
Estimation and Valuation of Benefits and Costs	350
<i>Can We Put a Price Tag on the Intangibles of Life?</i>	351
<i>How Can We Measure Future Loss or Gain?</i>	353
Introduction of Time: Present Value Analysis	355
<i>Choice of Time Horizon</i>	359
<i>Choice of Discount Rate</i>	360
<i>The Internal Rate of Return</i>	362
Choosing the Best Alternative	364
<i>Sensitivity Analysis</i>	364
<i>Cost-Benefit Ratio</i>	364
<i>The Limits of Cost-Benefit Analysis: Redistribution of Income</i>	366
<i>Cost-Effectiveness Analysis</i>	366
Key Words	369
Exercises	369
Notes	370

CHAPTER 15 SO YOU WANT TO BE AN ANALYST? SOME PRACTICAL SUGGESTIONS 373

Before You Start	374
<i>"Know Thyself" and Your Organization</i>	374
<i>Be Aware of Biases</i>	375
<i>Be Aware of External Constraints</i>	376
Begin Your Analysis	377
<i>Organize Your Thoughts: A Quick Analysis</i>	377
<i>Conducting a More Elaborate Analysis</i>	378
Tell a Good Story: Effective Presentation	382
<i>Know Your Audience</i>	383

<i>Know What You Don't Know</i>	383
<i>Say It with Numbers</i>	384
<i>Tell a Compelling Story</i>	385
Whose Ball Is It Anyway? Zen and the Art of Public Policy Analysis	387
Exercises	387
Notes	387
Appendix A Example of a Policy Analysis Report	389
Appendix B Models of Causal Prediction: Multiple Regression	412
Appendix C Areas of the Standard Normal Distribution (the Z table)	447
Appendix D Critical Values of the <i>t</i> Distribution	448
Appendix E Critical Values of the <i>F</i> Statistic: 5 Percent Level of Significance	449
Appendix F Critical Values of the <i>F</i> Statistic: 1 Percent Level of Significance	450
Appendix G The Chi-square Distribution	451
Index	453

Tables, Figures, and “A Case in Point”

TABLES

1.1	Ranking of Nations Based on Academic Performance of 15-Year-Olds	4
1.2	Classification of Goods	6
3.1	Voters' Preferences on the Decision to Purchase Fire Engines	64
3.2	Voters' Preferences on the Decision to Purchase Fire Engines: Strategy Change by Group A	65
4.1	Comparison of Pluralist and Elitist Models of Agenda Setting	83
4.2	Outcomes for Diffuse and Specific Costs and Benefits	86
5.1	Multiple Goals: Building a New Stadium	105
5.2	Correlation Analysis: Marital Status versus Student Achievement	114
6.1	Distribution of Housing Prices	133
6.2	Weighted Observations	136
6.3	Grouped Data	137
6.4	Relative Distribution of Housing Prices between Two Neighborhoods	138
6.5	Relative Distribution of Housing Prices: Deviation from the Mean	139
6.6	Calculation of Variance	140
6.7	Comparing the Dispersion of Two Distributions	141
6.8	Data Showing a Symmetric Distribution	142
6.9	Examples of Symmetric and Asymmetric Distributions	143
6.10	Mean, Median, and Mode in Symmetric and Asymmetric Series	144
6.11	Stem-Leaf Analysis: Drug Arrests, $N = 16$	145
6.12	Crime Rate and City Size	147
7.1	Probability and Cumulative Probability Distribution of Caseloads for a Social Worker on a Given Day	153
7.2	Comparison of Means	164
7.3	Example of a Contingency Table	166

7.4	Errors of Hypothesis Testing	168
7.5	Risk Assigned to Parolees	170
7.6	Quadratic Score Matrix for Parolees	170
9.1	State and Federal Grants to Masters, Pennsylvania, in Current Dollars, 2000–2009	204
9.2	State and Federal Grants to Masters, Pennsylvania, in Current and Constant Dollars, 2000–2009	204
9.3	Consumer Price Index (CPI): All Urban Consumers, 1990–2008	206
9.4	State and Federal Grants to Masters, Pennsylvania, as a Percentage of Revenue, 2000–2009	208
9.5	Yearly Percentage Change in State and Federal Grants to Masters, Pennsylvania, in Constant Dollars, 2000–2009	208
9.6	Index of State and Federal Grants to Masters, Pennsylvania, 2000–2009	210
9.7	Spending by Function and Level of Government, 1995	222
10.1	Projected Demographic Profile of New Community	229
10.2	Relative Issue Positions	238
10.3	Availability of Resources	239
10.4	Potential for Policy Influence	240
10.5	Ranking of Resources	240
10.6	Calculated Policy Influence	241
11.1	Tourist Population of Masters, Pennsylvania	257
11.2	Calculation of Quarterly Adjustment Factors	258
11.3	Seasonally Adjusted Quarterly Tourist Population	258
11.4	Calculation of Yearly Adjustment Factors	260
11.5	Trend-adjusted Quarterly Tourist Population	260
11.6	Quarterly Series on the Number of Tourists Visiting Masters, Pennsylvania	262
12.1	Population over Time	277
12.2	Calculations for a Straight-line Regression Equation	278
12.3	Calculation of R^2	280
12.4	Calculations for Tests of Significance	286
12.5	Per Capita Federal Grant Expenditures	288
12.6	Setting Up an Intercept Dummy Variable	290
12.7	Expenditures on Drug Rehabilitation Programs	292
12.8	Setting Up Slope Dummy Variables	294
12.9	Comparison of Estimated Data	296
12.10	Budget Deficit as a Percentage of GDP (five-year average)	297
12.11	Transformation of Data for the Estimation of a Quadratic Equation Form	298
12.12	Yearly Crime Statistics	312
13.1	Hypothetical Expected Payoff Matrix for President Ford's Response to Possible Swine Flu Epidemic	323
13.2	Estimated Flood Damage with and without Flood Wall	325

13.3	Estimated Flood Damage with Flood Wall and Levee	327
13.4	Total Estimated Cost of Flood Wall and Levee	330
13.5	Payoff Matrix for Mayor and Garbage Collectors' Union	334
13.6	Possible Days of Allied Bombing	336
13.7	The Prisoner's Dilemma	338
14.1	Benefits and Costs of Convention Center	349
14.2	Comparison of Costs and Benefits for Projects A and B	357
14.3	Contradictory Decision Based on Net Difference and Benefit-Cost Ratio	364
14.4	Costs and Benefits of Library Facilities	365
14.5	Costs and Benefits of Library Facilities, Ranked According to Benefit-Cost Ratio	365
15.1	Steps in Conducting an Effective Analysis	374

FIGURES

1.1	Goals and Their Trade-offs	15
1.2	World Military Spending	22
2.1	Four Models of Policymaking	31
2.2	Alternative Responses to Value Conflict	34
3.1	The Foundation of Economic Reasoning: Market Equilibrium	48
3.2	Marginal Utility and Consumer Surplus	50
3.3	Rent Seeking	57
4.1	The Policy Cycle	76
4.2	The Pluralist Model of Agenda Setting	77
4.3	The Elitist Model of Agenda Setting	81
4.4	The Subgovernmental Model of Agenda Setting	82
4.5	Decision and Nondecision in Policy Formulation	90
4.6	Impact of External Shocks on Policy Coalitions	93
5.1	The Pareto Principle	107
5.2	Methodology for Policy Analysis	110
5.3	Solomon Four-Group Experimental Design	111
5.4	Post-test-only Control Group Design	112
6.1	Class Divisions in Presidential Voting, 1952–2004	132
6.2	A Symmetric Distribution	134
6.3	Distribution with an Extreme Value	135
6.4	Distribution with the Modal Value	136
6.5	The Effects of Different Dispersions	138
6.6	The Range, the Midspread, and Hinges	139
6.7	Example of a Symmetric Distribution	142
6.8	Comparison of Symmetric and Asymmetric Series	143
6.9	Stem-Leaf Method and Histogram	145

6.10	The Extent of Correlation between Two Variables	146
6.11	The Nature of Correlation	146
7.1	Probability and the Number of Tries	151
7.2	Plot of Probability Distribution	154
7.3	Normal Distributions with Different Means	154
7.4	Normal Distributions with Various Levels of Dispersion	155
7.5	Standard Normal Distribution	155
7.6	One-tailed Test	159
7.7	Two-tailed Test and Confidence Interval	161
8.1	Measuring the English Shoreline	179
8.2	Errors in Sampling	181
8.3	Random Sampling Errors	185
8.4	Searching with the Computer	194
9.1	State and Federal Grants to Masters, Pennsylvania, in Current Dollars, 2000–2009	205
9.2	State and Federal Grants to Masters, Pennsylvania, in Current and Constant Dollars, 2000–2009	207
9.3	State and Federal Grants to Masters, Pennsylvania, as a Percentage of Local Government Tax Revenue, 2000–2009	209
9.4	Yearly Percentage Change in State and Federal Grants to Masters, Pennsylvania, in Constant Dollars, 2000–2009	210
9.5	Yearly Percentage Change in State and Federal Grants to Masters, Pennsylvania, by Indexing, 2000–2009	211
9.6	Expenditures on a Drug Prevention Program in Masters, Pennsylvania	212
9.7	Ratio of Nonwhite to White Unemployment Rate	218
9.8	Ratio of Nonwhite to White Unemployment Rate (stretched graph)	219
9.9	Ratio of Nonwhite to White Unemployment Rate (elongated vertical axis)	219
9.10	Ratio of Nonwhite to White Unemployment Rate (1960, 1970, and 1980)	220
9.11	Ratio of Nonwhite to White Unemployment Rate (1960, 1970, and 1980; shortened vertical axis)	220
9.12	African Elephant Population, 1980–1990	221
10.1	Trend Projection versus Causal Prediction	226
10.2	Problems of Single-Factor Projection	231
10.3	Relative Issue Positions on Public Funding of AIDS Information	239
10.4	Expected Payoff of the Government and Its Opposition	243
10.5	Dynamics of the Policy Outcome	244
11.1	Number of Tourists in Masters, Pennsylvania	250
11.2	Linear Trend Patterns	251
11.3	Horizontal, or No-Trend, Pattern	251
11.4	Seasonal Trend Pattern	252

11.5	Effects of Economic Cycles	252
11.6	Discerning Trend Patterns in the Tourist Data from Masters, Pennsylvania	253
11.7	Quadratic Trends	254
11.8	Exponential Trends	255
11.9	Logistic Trend	255
11.10	Catastrophic Trend	256
11.11	Plot of Number of Tourists per Quarter	257
11.12	Plot of Seasonally Adjusted Quarterly Data	259
11.13	Plot of Trend-adjusted Tourist Data	261
11.14	Adjusted and Unadjusted Data on Tourists Visiting Masters, Pennsylvania	264
12.1	U.S. Gross Domestic Product per Capita (in 1972 constant dollars)	272
12.2	Time-series Model	273
12.3	Errors of Estimation	274
12.4	The Problems of Summing the Errors	275
12.5	Explanation of a Straight-line Equation	276
12.6	Population Trend	279
12.7	Curve of Normal Distribution	281
12.8	Increasing Accuracy of Estimated Coefficients	287
12.9	Abrupt Shift in a Series	289
12.10	Shift in Expenditure Pattern	290
12.11	Plot of Trend in Government Expenditures for Drug Rehabilitation	293
12.12	Comparison of Estimated Trend Lines	295
12.13	Publicly Held Federal Debt as a Percentage of GDP: Linear Estimation	298
12.14	Publicly Held Federal Debt as a Percentage of GDP: Quadratic Estimation	299
12.15	Prediction with a Third-degree Polynomial	300
12.16	Symmetry Forced by a Quadratic Form	300
12.17	Double-Log Functional Forms	302
12.18	The Inverse Form	303
12.19	Errors of Prediction	310
13.1	Hypothetical Decision Tree for Swine Flu Threat	323
13.2	Decision Tree for Swine Flu	324
13.3	Probability of Flood Damage	326
13.4	Decision Tree for Flood Control	326
13.5	Expanded Decision Tree for Flood Control with Levee Option	328
13.6	Decision Tree for Flood Control with Total Expected Costs	329
13.7	Risk and Payoff in a Television Game Show	331
13.8	Strategies in the Battle of the Bismarck Sea	335
13.9	Decision Tree for Allied Strategy in the Battle of the Bismarck Sea	336
14.1	Classification of Costs and Benefits	348
14.2	Consumer Surplus and Social Benefits	350
14.3	Time Preference and Discount Rate	357

14.4	Plot of Present Values as a Function of Discount Rates	359
14.5	The Effects of Choosing a Time Horizon	360
14.6	The Internal Rate of Return	363

"A CASE IN POINT"

Moral Hazard and the Banking Crisis	10
What Do You Mean by Government Efficiency, Equity, and Liberty?	17
The Numb and Number: The Exploding Domain of Public Policy Analysts	29
Ignored Advice and the Global War on Terror	33
The Saboteur, the Whistleblower, and the Loyal Protester	36
Poverty in America: Contrasting Perspectives	40
Too Big to Fail: How Monopolies Can Hold the Government Hostage	54
Drug Policy: Our Addiction to Failure?	58
College Affordability: Should Government Intervene?	60
Trading Pollution: Simulating the Market	71
What's in a Name? The Case of the Mexican Swine Flu Pandemic	78
The Power of a Focus Event: Making Schools Safe	84
Halliburton and Government Patronage: Rent Seeking during the War	118
Do You Give a Dam? The Fate of the Snail Darter	121
Probability in Baseball: Revenge of the Nerds	152
Errors of Observation: The Hawthorne Experiment	183
Focus Group for Growth Management Policies	191
Politics and Government Statistics	197
The Politics of Data Collection: The Case of Terrorism Research	198
Racial Profiling	212
Digging Deep into the Numbers: A Stolen Election?	216
Single-Factor Analysis: The Fiscal Impact of Illegal Immigration in San Diego County	231
The Use of the Delphi Technique in Devising a Public Policy Curriculum	236
Fighting over a Fiction? The Budget Surplus in the Presidential Debate	266
Campaign Contributions and Judicial Impartiality: How Regression Analysis Can Help	282
Who Wins Olympic Medals? Beijing 2008	308
Risky Business: Orange County Bankruptcy	332
Fitting Seat Belts in Texas School Buses	367

Reason, Rationality, and Public Policy: The Puzzle of Human Behavior

GOVERNMENT IN OUR LIVES

In "Tradition," the opening number of the hit Broadway musical turned movie *Fiddler on the Roof*, Tevye the milkman introduces the rabbi as the most important person in the tiny Jewish village of Anatevka: "[The rabbi] tells us what to eat, when to eat, how to dress, when to pray." The audience wonders about this distant land and its alien customs. Life in this nineteenth-century village seems so different from modern life in our own country; here, no one tells us how to live. Or do they? Upon reflection, we realize that an outside force shapes practically every aspect of our lives. In the modern world the government regulates almost everything we breathe, touch, use, ride, inject, and ingest, from our birth to our death. The government sets air and water quality standards and criteria for serving food and administering drugs. It levies taxes, circulates currency, makes education policy, and resolves questions of our personal welfare. The government tries to protect us by keeping law and order. It maintains a military to keep us safe from attack. It even passes laws affecting the most personal aspects of our lives, including laws regulating sexual behavior. State governments determine the age of consent, prohibit or permit same-sex marriages, and define acceptable sexual practices. It would be difficult, if not impossible, to list all the government regulations and standards to which we must adhere in our daily lives.

Consider, for example, the birth of a baby. The hospital must follow strict rules and regulations governing medical equipment, medicines, and medical personnel. The parents must register the child's birth. When they leave the hospital and get into the family car, they cannot simply place the newborn on the mother's lap. In many states, the infant must be strapped into an approved car seat. Similarly, when we die, many government regulations tell us how our family may dispose of our remains. Imagine how long the list would be if you

wrote down all the ways government policies have touched your life since you got out of bed this morning.

The government does not make public policies in a vacuum. Its control over its citizens depends on several factors that determine what are and what are not acceptable domains of public policy. These factors include the nation's norms, values, culture, history, traditions, constitution, and technological sophistication. Together, these factors form the environment in which public policies are made. In the United States and other Western nations, the rights of the individual reign supreme; citizens challenge immediately any policy that impinges on these rights. Therefore, policymakers seek to uphold individual rights. In contrast, in collective societies ruled by national religions or communism, lawmakers value the needs of the group more highly than the rights of individuals. In making public policy, such societies seek to uphold not the rights but the duties of citizens. Some countries arrest those who do not pray during specified times of public prayer, or impose fines for violating the Sabbath. Other countries severely curtail the right of political expression. Although most Western countries consider property rights sacred, other nations balance these rights against competing national goals, such as eliminating poverty or meeting society's general needs, as defined by the political elite.

The extent to which government intrudes in the lives of its citizens is partly determined by technology. As technology advances, so does society's need for regulations. A simple agrarian society requires minimal government intervention. However, today's complex world, characterized by rapid changes in technology—from the Internet to genetically altered food, from the possibility of cloning animals and humans to the specter of global warming caused by unrestrained human activities—demands regulations with increasing urgency.

In the wide-open field of public policy analysis, the analyst performs a high-stakes balancing act. Like the fiddler perched precariously on a slippery roof, playing his fiddle, public policy analysts ply their trade by balancing a number of competing demands, confusing perceptions, and other pitfalls of scientific reasoning.

POLITICS AND RATIONALITY

Politics is sometimes defined as deciding "who gets what, when, and how." In a democratic free market system, government allocates scarce resources among competing demands. In giving someone or some group something, government often has to take it from someone else. As a result, society is always rife with conflict. The study of politics allows us to understand the process by which actors receive government-allocated public goods, while the study of *public policy* lets us understand the rationale for the allocation. In other words, it adds to the four questions—*who*, *what*, *when*, and *how*—the fifth important question of *why*. Policy analysts provide the rationale and draw up the roadmap for the allocation of goods, advising the decision makers—elected public officials—regarding the proper course of action. In order to understand the rationale of allocation, however, we must begin with an understanding of **rationality**.

Do you consider yourself to be a rational individual? Of course you do. Few of us would admit to being otherwise. Although we take it for granted and use it in our everyday parlance, what does the term *rational* really mean? To most of us, being "rational"

means analyzing a situation based on logic, without any interference from emotion or instincts. Separation of emotion from reason has been the cornerstone of Western philosophy. Plato, René Descartes, and Immanuel Kant all proclaimed that what separates humanity from all other species on Earth is our ability to overcome natural instincts and emotion. Following this tradition, Adam Smith, the father of modern economics, in his 1776 book *The Wealth of Nations*, assumed that everyone is a “rational” actor. His assumption was based on three principles: (1) each of us is aware of what is best for us; (2) our behavior is dictated by the simple principle that we act when the benefit of an action is greater than the corresponding cost; and (3) among various alternatives, we choose the one that gives us the most benefit compared with the cost. Smith argued, “It always is and must be the interest of the great body of the people to buy whatever they want of those who sell it cheapest. The proposition is so very manifest, that it seems ridiculous to take any pain to prove it.”¹

Thus, Smith rhetorically asked, if you are in a shop to buy a piece of cloth and one is more expensive than the other, will you spend more money to get an identical product? Obviously not. If you are a seller and you ask for more money for your product than your competitors, nobody will buy from you. On the other hand, if you are offering a price that is less than market value, the demand for your product will be so overwhelming that you will be forced to raise your prices. (For a fuller explanation of this supply-and-demand process, see chapter 3.) In a free market, therefore, nobody can get more than what the market will bear. To Adam Smith, the market process was not only inherently the most efficient but also the fairest: nobody will be able to get more or less than what he or she deserves. This market process is rational, since you don’t make your market decisions according to emotions; rather, you buy or sell according to the market price. You may cooperate with another person, but only on the basis of your self-interest. When we all behave in our best individual interest, the market achieves its greatest point of efficiency and reaches the pinnacle of distributive justice. Adam Smith called this process the **invisible hand**, which solves all market-related problems through the working of an unseen force, greater than the wisdom of any one individual or group. This force is always impersonal, totally impartial, often cruel, but, in the end, always fair. Therefore, the best course of action for the government is to interfere as little as possible in the natural course of the market, only providing internal security through maintaining the rule of law and protecting the nation from external threats.

Unfortunately, however, as we will see in the following chapters, the market does not always behave “efficiently.” The shortcomings of the market were clearly exposed during the Great Depression of the 1930s. Economists, starting with John Maynard Keynes, pointed out that at certain times the market fails to correct itself, and the government must step in to remedy market inefficiency through regulation, restriction, or direct infusion of money. Keynes’s economic theories shaped our economic policies during the Great Depression and laid the foundation for the Obama administration’s efforts to stimulate the lagging economy in 2009.

In the policy world, the identification of a social need, caused by market inefficiency, usually starts in the political arena. Let us take a hypothetical example to illustrate this process of rational decision making, which is also often called the **policy cycle**. We will have an opportunity to examine this policy cycle in greater detail in chapter 4.

TABLE 1.1**Ranking of Nations Based on Academic Performance of 15-Year-Olds**

Math (2006)	Reading (2003)
1. Finland	1. Finland
2. South Korea	2. South Korea
3. The Netherlands	3. Canada
4. Switzerland	4. Australia
5. Canada	5. New Zealand
6. Japan	6. Ireland
7. New Zealand	7. Sweden
8. Belgium	8. The Netherlands
9. Australia	9. Belgium
10. Denmark	10. Norway
11. Czech Republic	11. Switzerland
12. Iceland	12. Japan
13. Austria	13. Poland
14. Germany	14. France
15. Sweden	15. United States
16. Ireland	16. Denmark
17. France	17. Iceland
18. United Kingdom	18. Germany
19. Poland	19. Austria
20. Slovak Republic	20. Czech Republic
21. Hungary	21. Hungary
22. Luxembourg	22. Spain
23. Norway	23. Luxembourg
24. Spain	24. Portugal
25. United States	25. Italy

Source: National Governors Association, Council of Chief State School Officers, and Achieve, Inc., "Benchmarking for Success: Ensuring U.S. Students Receive a World-class Education," 2008, www.nga.org/Files/pdf/08128BENCHMARKING.PDF. © National Governors Association. Reproduced by permission of the National Governors Association. Further reproduction prohibited.

It is an easily demonstrable fact that education is the key to economic success. Yet as Table 1.1 shows, the standardized test scores of high school students in the United States are among the lowest compared with other developed countries. Since the academic preparation of our youth determines the future prosperity of the country, let us suppose that this has prompted the president of the United States to ask the Department of Education (ED) to develop appropriate policies to increase the scores by a certain percentage within a specific time frame. The policy analysts in the ED accept the task and proceed to formulate appropriate policy prescriptions.

Policy analysts begin work by observing the real world and collecting relevant data to measure where we are at this particular point in time and where we need to be at the end of the policy period. Therefore, the policy formulation process begins by collecting data on current student performance. Once armed with those data, the analysts will consider alternative plans to bridge this performance gap, since there is no single "correct" method for achieving the policy goal. These **alternatives** might include strategies such as providing additional funds for hiring qualified teachers, appointing educators to improve the curriculum, increasing the time students spend in school, emphasizing

charter schools, starting after-school tutoring programs, working with parents to instill the value of education in their children, or even providing greater assistance to school lunch programs so that poor children are not distracted by hunger.²

The next step toward a scientific analysis is to develop a **model**. A model is a scaled version of complex reality that allows us to test our hypotheses regarding cause and effect. The ED analysts will take each alternative and project how it will improve test scores—we will discuss these methods of projection at length in later chapters. The foundation of scientific reasoning lies in our ability to test the validity of hypotheses with empirical evidence. The eminent philosopher of science Karl Popper called this **falsifiability**. For example, the hypothesis that hiring better prepared teachers is the most cost-effective way of improving students' performance is falsifiable—it can be proven false with contrary evidence. In contrast, questions regarding the existence of God, the soul, or spirits, which cannot be observed, measured, or tested, are essentially beyond the scope of scientific inquiry.

Based on the projected costs and benefits produced by the models, the ED analysts will make their recommendations to the president, who will propose a course of action. Once legislators are able to pass a bill on the action, and the president signs it into law, money starts flowing from the treasury to fund the new policy. This is the beginning of the implementation stage of the policy cycle. After the policy is formulated, legislated, adopted, and implemented, the analysts evaluate the results against evidence from the real world. Did the policy achieve its goals? What did we learn from our experience? In policy analysis, this step is known as evaluation research.

This process sounds simple and makes sense. If we were all "rational" beings, in the strict economic sense of the term, social concerns would be readily validated by data, models would produce irrefutable projections of costs and benefits, and the most efficient policy would be adopted and implemented without any dissent. In short, we would achieve our policy goals. Unfortunately, as we will see, at each stage of the process our lives become complicated as the precepts of "rationality" are violated. Let us now see why this eminently "reasonable" description of policy analysis becomes convoluted in actual life.

THE WORLD OF THE "RATIONAL FOOL"

Adam Smith took it for granted that we are all eminently rational.³ And why not? Individuals are selfish—they are motivated by self-interest. This assumption of human nature permeates political discourse and is enshrined in our Constitution. It is so commonplace that we are barely conscious of this underlying assumption of human nature while discussing public policies.⁴ Yet how valid is such an assumption? Political scientist George Lakoff points out, "Most of us have inherited a theory of mind dating back at least to the Enlightenment, namely, that reason is conscious, literal, logical, unemotional, disembodied, universal, and functions to serve our interests. This theory of human reason has been shown to be false in every particular, but it persists."⁵ Since public policies aim at altering individual behavior, their formulation requires us to understand the roots of motivation; we need to know what incentive will nudge people to act in a certain way. Thus, the Nobel laureate economist Amartya K. Sen, in his influential essay "Rational Fools," points out that "universal selfishness as *actuality* may be false, but universal selfishness as a requirement of *rationality* is patently absurd."⁶

Over the past few decades, research from fields as diverse as biology, cognitive science, primate behavior, social psychology, and even experimental economics has painted a much more complex picture of this intensely social being. In this chapter, I will argue that the two interrelated obstacles to the rational model of policy analysis arise out of our own nature and the society we have created as a consequence. Specifically, I will show that

- the assumption of a self-utility-maximizing being leads us to a logical contradiction, particularly when it comes to the provision of public goods, the core of public policy analysis;
- as social and emotional beings, we often understand the world in profoundly different ways; and
- group identities are extremely important in the understanding of the public policy process.

The Illogic of a “Rational Being”

The concept of a perennially self-utility-maximizing human being, the fundamental building block of economics, presents a strong logical contradiction when it comes to the provision of public goods, the core subject matter of this book. Before I present my arguments, let me first define the term **public good**.

We can classify all goods with the help of two characteristics, **excludability** and **exhaustibility**, which is also called the **rivalrous** quality of a good (see Table 1.2). If your friend asks to borrow this book from you, you can say no, since this is your own copy and you have every right to deny access to it (although you may not have a friend much longer). In contrast, if the book belongs to the library, after your time is up, your friend has an equal right to borrow it. The difference between these two cases lies in the book’s excludability; a user can exclude others from using certain types of goods. Exhaustibility, on the other hand, depends on whether the utility of a good to its user goes down when an additional user joins in. In other words, is there any rivalry in the good’s use? For example, if you are driving down the highway, your ability to use the highway does not go down simply because one more car is on the road (the way it does when another person wants to read your book at the same time); thus, the utility of the highway is nonexhaustible. I have provided a matrix in Table 1.2 to classify four types of goods that we enjoy in life.

Let us go back to the example of the book you are reading. If you own it, it is purely a **private good**. You can limit access to it, since sharing it with someone else will decrease its utility to you. You may also decide to buy the book jointly with a few other friends.

TABLE 1.2

Classification of Goods

	Exhaustible (rivalrous)	Nonexhaustible (nonrivalrous)
Excludable	Private goods	Toll goods
Nonexcludable	Club goods (common pooled resources)	Pure public goods

In that case, although it is essentially a private good within your friend circle, you decide to use it as a public good. Everyone within this small group has an equal right to it. In economics, this is known as a **club good** (or, alternatively, a **common pooled resource**), because the benefits of owning it remain within the group, or the "club." On the other hand, there are some goods that are actually meant to be shared, such as a bridge, a road, public transportation, or a national park. Yet because of the possibility of overuse when they are provided free of charge, or because they need to be financed, a toll may be required in order to use them. The fee on a **toll good** restricts its use only to those who are willing to pay and excludes those who are unable or unwilling to do so. Finally, we are left with the category of **pure public goods**. These are goods that must be shared with everyone, even those who cannot or will not pay for their use. Suppose you live in Miami Beach, Florida, a town known for its beautiful beaches. It costs the city (and the local taxpayers) money to provide lifeguards at the public beaches. If I am visiting Miami Beach and, therefore, do not pay city taxes, I am still fully eligible to swim in the ocean under the watchful eyes of the lifeguards, and should I need their services, they will be provided to me equally as to any other local taxpaying citizen.

Among these four types of goods, the question of the allocation of pure public goods is most problematic. For the other three types of goods, there is a strong link between benefits and costs. When you pay for a private good, club good, or toll good, you know what you are getting in return for your money. As a result, consumers are able to adjust their demand according to price. However, when it comes to public goods, the link between benefits received and the price paid is obscure. For example, we pay for clean air, a public good, through our tax dollars that are spent on monitoring the levels of pollutants in the air. We also pay for it through higher prices when, say, a tax is imposed on gasoline, or the government directs oil companies to use less polluting but more expensive additives or forces car manufacturers to install hybrid engines. As we pay directly or indirectly for the quality of the air, we have no control over "how much" clean air we "purchase." Some of us might want cleaner air and be willing to pay a higher price for it, but others would rather keep their costs down. Yet when the government sets air pollution standards, one size must fit all. On top of these variations, there are many poor people or visitors from other nations who also enjoy a clean environment without paying anything at all for it. In other words, the adjusting market mechanism of costs and benefits, which functions for the other three types of goods, does not operate for public goods. As a result, there are constant debates on the best use of government money and the allocation of these public goods.

The aftermath of World War II and the onset of the Cold War era represented a "feel good" time in public policy literature. The contrast between a free market democracy and communism was stark. An overarching sense of triumph was pervasive as American scholars of public policy often predicted the eventual victory of democratic values over the twin evils of dictatorship and communism. It was assumed that a free society, in which every person was dictated to by nothing other than his or her self-interest, would find ways of allocating government-provided goods efficiently and ethically. For example, if there was a need or inefficiency, the citizens would band together, lobby the government, and correct the inefficiency. Suppose the folks in your community noticed that a busy street intersection needed a traffic light. They would immediately form a neighborhood committee, work through the democratic processes, and get the local

government to install a traffic light. If the authorities were unresponsive to the popular demand, they would risk getting booted out of office in the next election. How different, people wondered, were lives in a communist or a nondemocratic country, where common people were deprived of their ability to mobilize for collective action.

However, in the late 1960s influential scholar Mancur Olson created a stir by showing that when it comes to the allocation of public goods through individual effort, the assumption of the self-utility-maximizing economic being, or *homo economicus*, runs us into a logical quagmire. Simply put, the problem for Olson was that since the benefits of a public good are not dependent on an individual's effort, why should a rational person—figuring out that she can get something for nothing—contribute toward its procurement? As I drive to school, I listen to National Public Radio (NPR) without paying anything, just as I can to any other station. NPR differs from other radio stations, however, in that a large part of its funding comes from its listeners' contributions. As I contemplate contributing money, I realize that even if I don't contribute a single penny, others will, and I will continue to get something for nothing. Second, I also know that my contribution of "a dollar a day," which is a significant amount in my own finances, will make only a marginal difference in the survival of my favorite station. Taken together, does it make any *economic sense* for me to send in my pledge? Absolutely not. If we all act rationally, no one will choose to contribute to NPR. And as a result, no public good—from a neighborhood streetlight to measures against global warming—would ever be produced. Olson called this dilemma the **free-rider problem**, and in political science, it is also known as the **collective action problem**.⁷

Mancur Olson's work paved the path for brand new ways of looking at society, especially for those who want to know how to make the government more efficient. Among many interrelated issues, we can talk about a few. First, let us consider **moral hazard**.⁸ Suppose you know that your insurance policy covers car theft. In that case, what is your incentive to take the extra step to make sure to lock your car doors? To extend this example, when Congress bails out yet another very large bank or an automobile manufacturer, what is the incentive for those receiving bailout money to run an efficient shop if they know that their own risks for bad decisions are close to zero (see "A Case in Point")? What is the guarantee that in a large workplace, some workers will not be tempted to **shirk** their responsibilities, knowing that nobody would find out that they are not pulling their weight?⁹ These questions illustrate the problem of moral hazard.

Consider another class of problems. If I make a commitment to you, and there is no oversight and I face no cost for going back on my word, what is my incentive to keep my word when it does not suit my purpose to do so? In other words, if my commitment is not credible, why should anyone trust me? In the literature, this is known as the **credible commitment problem**.¹⁰ For instance, the United States is extremely concerned that Iran might build its own atomic weapon, which might destabilize the entire region. Therefore, in order to dissuade Iran from becoming a nuclear power, the United States threatens "severe consequences" whenever Iran makes a move that indicates its progress toward manufacturing nuclear weapons. Yet in the international arena, particularly facing opposition from Russia and China, the hands of U.S. foreign policymakers are tied. Therefore, without a credible commitment, Iran may soon develop the dreaded weapons.

Finally, the free-rider problem can lead to the **tragedy of the commons**.¹¹ Imagine a town in which large plots of public land are set aside for grazing cattle. If everybody in the village lets their cattle feed on the grass for a set amount of time, the grass has time to regenerate, and everyone benefits from the continuing use of the land. However, if I keep my cattle on the land a bit longer, then my better fed cattle are fatter, and they fetch a higher price in the market. But if everybody lets their cattle feed for longer hours (since we are all self-utility-maximizing *homo economicus*), we will soon see an environmental catastrophe in which the grass dies out and everybody suffers. Viewed from a different angle, what is my incentive to abide by the rule? If I limit the time my cows are grazing on the public land, but if I am the only one doing so, my noble self-restraint does little to improve the overall condition of the grazing land. Such a situation can be extremely real and pose huge policy problems.

In his book *Collapse*, biologist Jared Diamond shows how the unrestrained pursuit of self-interest brought about an environmental collapse on Easter Island in the Pacific Ocean. When tribes started competing with each other for natural resources on the isolated island, they irrevocably damaged the fragile environment, which did not have the opportunity to regenerate, causing mass starvation, wars, and the death of most Easter Islanders. We can also see the same scenario being played out in our world as unchecked commercial fishing depletes stocks all over the ocean; homeowners do not conserve groundwater and overwater their lawns, which lowers the groundwater to a level where it is nearly impossible to access; or the industries that belch pollutants from their smokestacks help create acid rain that kills many living things in the vicinity. In short, the free-rider problem points to our irresponsible behavior in pursuit of short-term, selfish utility, as well as the need for suitable regulatory action by authorities who are mindful of the welfare of the entire community (a public good).

The “Un-rational Being” and Political Behavior

Insightful as Mancur Olson's arguments are, the timing of his book's publication, in 1965, was a bit ironic. Just as Olson was pointing out why “rational actors” would not participate in a collective action, tumultuous mass movements were occurring in the United States. The 1960s saw the development of the civil rights movement, joined by people of every color who were inspired by Dr. Martin Luther King Jr. and prodded by the sight of racial intolerance in the South. There were also violent movements, such as the Weather Underground organization, an anarchist group that wanted to bring down organized society and its manifest inequalities, and the Black Panthers and the Symbionese Liberation Army, both of which wanted to end segregation through violent means. The members of these groups did not think that their individual actions would bring about social change, and they knew very well that they were risking pain, prison sentences, or even death. Regardless of the validity of their vision, each member set aside his or her personal welfare in order to seek a certain public good, to change the nation in their way. Olson was not oblivious to these developments and thought that many human actions were not based on “rational” calculations. He pointed out, “It is not clear that this is the best way of theorizing about either utopian or religious groups. . . . Where nonrational or irrational behavior is the basis for a lobby, it would perhaps be better to turn to psychology or social psychology than to economics for a



A CASE IN POINT

Moral Hazard and the Banking Crisis

In our economic system commercial banks play a central role by lending money to consumers to buy homes, cars, and other consumer goods, and to businesses for their expansion. If this money spigot is turned off, the entire economic system comes to a grinding halt, yet the banks cannot print money. They depend on us to open accounts and deposit money. Suppose I deposit \$100 in a checking account at my neighborhood bank. The bank knows that I will not come back tomorrow and withdraw all my money, so it keeps, say, 10 percent in reserve and lends out \$90 to someone who needs to borrow it. When the loan is approved, the bank manager does not come back with a satchel full of money but instead opens a checking account, showing a deposit of \$90. The process starts once again, when the bank keeps \$9 and lends out \$81. If you continue this calculation, you will see that at the end, the bank has created a total of nearly \$900 from my \$100 deposit. In effect, this is similar to a Ponzi scheme of borrowing from Peter to pay Paul. The banks make money when they lend money and the borrowers pay interest. However, if a significant number of borrowers fail to repay, the bank can fail. Therefore, banks have to be extraordinarily careful and resist the temptation to make risky loans.

The Great Depression began with the failure of a number of banks. As the news spread, people scrambled to get their money. The ensuing panic nearly brought down the entire monetary system. In order to prevent such a catastrophe, the Roosevelt administration created the Federal Deposit Insurance Corporation, which, by taking a small fee from all bank deposits, guaranteed that no individual depositor would lose money up to \$100,000 per account. (In 2009, the Obama administration increased this limit to \$250,000.) This move calmed the market but created a moral hazard for the banks. Since banks and bank executives make money when they lend, if

relevant theory. The beginning of such a theory may already exist in the concept of “mass movement.”¹²

Professor Olson was correct. Economists' method of analysis, as their critics point out, assumes that people are rational, predictable, and pretty much the same, regardless of culture, religion, or upbringing. In the four decades since the publication of Olson's book, research in diverse fields has shed new light on the mind of what I call the *un-rational* being. The flip side of the term *rational* is *irrational*, which implies a complete absence of reason or even insanity. We clearly are not so; we are simply not “rational” in the sense that the term is defined in traditional economics. It is important to note that this assumption about human nature plays a significant part in how we devise policies to deal with less desirable behavior.¹³ If the assumption regarding a person's motivations is not correct, it may also reflect in the policy that we prescribe to address a pressing social need. *New York Times* columnist David Brooks saw that a faulty assumption about human nature lay at the roots of the policy failure that led to the economic disaster of 2008 and 2009.¹⁴ He points out, “The nation essentially bet its future on economic models with primitive views of human behavior.” To Brooks,

there is no risk (and cost) to their bad decisions, what is the guarantee that they will not become too greedy and engage in overtly risky business practices?

In order to prevent this, Congress enacted the Glass-Steagall Act of 1933, which created a barrier between commercial and investment banks. The commercial banks were not allowed to invest in the stock and real estate markets, and the investment banks were barred from normal banking operations. The bank executives' excesses created such a public outrage that, a year later, Congress was able to create another important agency to oversee the powerful financial executives—the Securities and Exchange Commission. These measures worked well, but the banks chafed under the regulatory firewall and myriad rules governing their conduct.

The election of conservative president Ronald Reagan (who famously declared that government is not the solution but the problem) in 1980 opened the door for weakening the government's financial regulations. Over time, with the inclusion of banking industry insiders as heads of the agencies, the distinction between regulation and advocacy became blurred. Finally, in 1999, Congress passed and President Clinton signed into law the repeal of the Glass-Steagall Act, thus allowing commercial lenders to trade risky securities. Although there is no unanimity, there is a growing consensus among economists that the roots of the banking crisis of 2008 can be traced to the condition of moral hazard created by weak government oversight and a system that rewarded the banks for taking risks without much thought for the consequences of their actions.

Discussion Points

1. How is the notion of risk involved in the condition of moral hazard?
2. How does government temper the act of regulation in a free market system?
3. Does government intervention in the market amount to socialism?

the blame for the policy failure should be placed on the "policymakers and experts who based estimates of human psychology on mathematical equations." The complexity of the human mind stands in the way of objective scientific inquiry, particularly in the area of public policy. This complexity results from two interconnected concepts: the way social beings understand the world, and the way society influences policy research.

In the Western tradition, it is assumed that rational people keep their emotions separate from the cold calculations of benefits and costs. Cognitive scientist Antonio Damasio would beg to differ.¹⁵ Damasio cites the case of a man he calls Elliott, a successful business executive and a caring family man. After Elliott had a benign tumor removed from the frontal lobe of his brain, he successfully regained all his faculties, awareness, and speech, but he was a completely changed person in a profound way—he was unable to make a decision. Damasio discovered that the surgery had damaged a part of his brain tissue and permanently separated the emotional part of his brain from the analytical part. As a result, Elliott could perfectly analyze a complex situation, but he did not have the emotional backing to make the final decision.

While Damasio examined individual decision-making processes, Drew Westen, another prominent neuroscientist, studied how people make political decisions. During the heated presidential election between George W. Bush and John Kerry in 2004, Westen and his colleagues recruited fifteen diehard Democrats and fifteen diehard Republicans and scanned their brains in order to study how extreme partisans reason when they come face to face with information that directly challenges their strongly held beliefs—in other words, a contradiction between data (facts) and desire (political ideology).¹⁶ The scientists showed participants two slides: The first contained a statement that Bush or Kerry supposedly made (some were true, and others were fabricated), and the second showed another statement that to any reasonable person would seem contradictory. When the partisans saw the obvious contradictions, they reacted very similarly; the part of the brain that controls negative emotions fired up when Republicans saw contradictory statements made by Kerry, and Democrats' brains did the same when Bush's contradictory statements were shown. However, when it came time for their own candidates, a funny thing happened in the brains of both groups. At first a network of neurons that produces distress became active, indicating that the subjects were keenly aware of the contradiction and were distressed by it. However, the brain quickly "registers the conflict between data and desire and begins to turn off the spigot of unpleasant emotion. We know that the brain largely succeeded in this effort, as partisans mostly denied that they had perceived any conflict between their candidate's word and deeds."¹⁷ Even more amazingly, Westen found that "once partisans had found a way to reason to false conclusions, not only did neural circuits involved in negative emotions turn off, but circuits involved in positive emotions turned *on*."¹⁸ In other words, when it comes to politics and public policy, we most often make our decisions using our gut and political instincts, which may or may not involve questions of rationality or self-interest. Thus, emotion is not only legitimate but is also central to political persuasion.

As Westen's experiment demonstrates, people tend to seek affirmation of their core beliefs—especially through news sources that best fit their worldview—and this has resulted in polarization and a deepening of partisanship in the United States.¹⁹ A 2009 Pew Research Center poll shows that the gap in the public's perception of how the president is performing is getting wider with time. Traditionally, a newly elected president enjoys a "honeymoon" period, when Republicans and Democrats set aside their differences and give the incoming administration the benefit of the doubt. For example, a majority of Republicans actually approved of Jimmy Carter's job performance early in his term, and a majority of Democrats backed Richard Nixon at a similar point in his administration. A gap began to emerge when Reagan was elected, after he ran a highly ideological campaign. However, Barack Obama has the biggest partisan gap in his early job approval rating of any president in the past 40 years. The Pew poll shows that 88 percent of Democrats approve of the president, but only 27 percent of Republicans think he's doing a good job—a stunning 61-point gap that demonstrates the polarization between the collective identities of the two parties.²⁰

An examination of the "partisan brain" leads us to a discussion of the importance of group formation in our reasoning process. Our brains have evolved to gear us toward forming groups.²¹ For instance, social psychologist Henri Tajfel's experiments showed the universal human need to belong to a group and to strive for the group's welfare, even

when the act comes at the expense of one's own personal interest. Do you root for a professional sports team? If you do, does your mood depend on how "your team" performs on a specific day? If our teams do badly, some of us get mildly annoyed, some get angry, and some may even become violent, as English soccer fans in the 1980s and 1990s demonstrated.²² If you think about it, why should anyone be so emotionally wrapped up in what a group of millionaire athletes do on the field playing for their wealthy owners? In many instances, neither the players nor the owners have any permanent loyalty to any city; they often readily change teams or leave town when offered more money or better prospects. Yet so many of us remain loyal fans. Through our identification with a team (as opposed to a single player, such as Tiger Woods or Serena Williams), we exhibit our deep-rooted desire to belong to a group.

This need to have a group identity takes many different forms. Sociologist Benedict Anderson, for instance, points out that many of our cherished feelings, such as patriotism, are based on the creation of an "imagined community."²³ This preference for our groups or communities allows us to overcome the natural inertia of free riding; for example, young men and women—often spurred by patriotism or a sense of civic duty—join the armed services in times of war, knowing full well the risk of being sent to the front lines. We abide by the rules—both official and unofficial—of the "group" of society. Following our group norm, we dutifully lock our car doors even when we know that we are fully insured, we tip a certain percentage of the total bill, and we put in an honest day's work even when there are ample opportunities for shirking. Contrary to the rationality of *homo economicus*, we often behave altruistically toward others.²⁴ Most of these actions reflect our group behavior.

The formation of group identity is not accidental or random; it is often shaped by the deliberate acts of political leaders. In the political science literature, these leaders are referred to as **political entrepreneurs**, and they range from good to evil, from Gandhi, Lincoln, and Martin Luther King Jr. to Hitler, Stalin, and Osama bin Laden.²⁵ These political entrepreneurs often mobilize people by **framing** an issue. In psychological terms, a "framer" can change another person's perception of an issue by putting it in a certain context.²⁶ Take for instance the case of the federal inheritance tax. This tax was set for those who would inherit very large amounts of wealth and is known as the "estate tax." But in the 1990s, conservatives started calling it the "death tax." This new term was designed by a marketing expert to tap into the resentment that many feel toward government in general and taxes in particular. By changing its name, the Republican Party was able to mobilize public opinion, especially among the working class, against the tax.

Our group memberships lead to the formation of what social psychologists call collective identity—in other words, an "us" versus "them" mentality. When we separate ourselves into the categories of allies (the "in group") and enemies (the "out group"), we view the world through partisan glasses, allowing our group affiliations to dominate our rationale. Often this may prompt us to go against our self-interest. For instance, during his presidential campaign, Barack Obama advocated increasing taxes for those making over \$250,000 a year. Yet many affluent supporters of Mr. Obama rallied behind him and his economic agenda, even though his tax plan would disadvantage them personally. The fact of the matter is, when we view the world through the prisms of our group affiliations, we often violate the first principle of scientific inquiry: It is not that we

observe the world and understand; we often “understand” the world and then observe. If it were otherwise, by the fundamental precept of rationality, we would all come to the same conclusion when presented with the same set of facts. Even when there is no controversy over a fact, we tend to interpret it according to our political affiliations and ideological orientations (as Westen’s experiment demonstrated).²⁷

In the realm of public policy, group psychology explains consumer behavior during periods of inflation and recession. Take for instance the recent economic downturn, much of which was caused by troubles in the housing sector. The prosperity of the global economy fueled an increased demand for homes. As prices increased and stories of people making huge profits by buying homes and then selling them quickly began to circulate, everyone wanted to get in on the game. Fearing that they would be pushed out of the market, people lined up to buy homes in any way they could, often taking out loans (many of which seemed favorable in the short run, but became very expensive after the periods of “teaser rates” expired). With interest rates on mortgages at an all-time low, the risks of borrowing reduced, and an overabundance of available homes, many people threw caution to the wind—builders built too many houses, banks lent too much money, and homeowners bought too many properties. In short, the herd mentality took root. As the housing market took a sudden downward turn (or if homeowners lost their jobs), many homeowners suddenly faced bankruptcy and financial ruin, and it was inevitable that a large number of them would default on their loans. Indeed, Federal Reserve chair Ben Bernanke has singled out the herd mentality as an element that brought about the catastrophic swings in the market in 2008.²⁸

Finally, no discussion of public policy can be complete without noting that the fear factor dominates public discourse on policy changes. Every new policy introduces the prospect of change, and the uncertainty involved in change elicits fear and strong reactions, because of the human tendency to remember impending threats.²⁹ Fear is most often the primary motivator for collective action that produces the NIMBY (not in my backyard) factor, demonstrated most recently by the highly emotional debates surrounding comprehensive health care reform in the United States. Evolutionary biologists bolster the findings of **prospect theory** offered by Kahneman and Tversky,³⁰ which simply states that in the process of evaluating the benefits and costs of an action, human beings often place a far greater weight on the fear of a loss than the prospect of a gain. Jonathan Haidt points out, “If you were designing a fish, would you have it respond as strongly to opportunities as to threats? No way. The cost of missing the sign of a nearby predator, however, can be catastrophic. Game over, end of the line of those genes.”³¹ Any species that is fearless and does not take into account the prospect of a loss is destined to be extinct. Therefore, we must understand the essence of our primordial nature; in most policy debates, fear moves us in a profound way.

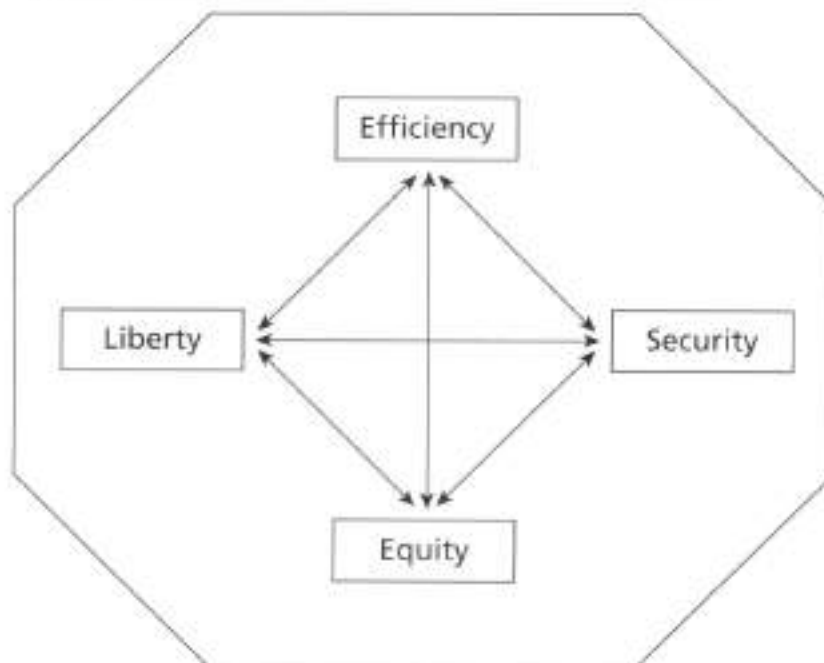
In sum, it is interesting to note that while most of us think of ourselves as “freethinking” individuals, our perceptions on issues become colored by many factors, including the groups in which we claim membership and the leaders and opinion makers to whom we listen.³² We also are guided by inherent moral and ethical concerns. If you recall my example of contributing to NPR, you should note that while there is no “rational” reason for me to contribute money, in reality I do, because I feel that contributing to the broadcast is the “right thing to do.” This simple appeal to our internal ethical code propels us to overcome our natural inertia to free ride.

GOVERNMENT'S GOALS AND THEIR IMPACT ON PUBLIC POLICY

If we draw out the concept of Adam Smith's "invisible hand," whereby competition in the free market process makes the government largely unwanted, we reach the ideal world of the libertarian philosophy. However, as we will see in chapter 3, the market process is not without its own shortcomings. Therefore, government intervenes in the market process to provide us with four desired goals: **efficiency, equity, liberty, and security**. Yet, as discussed above, these four cherished goals of public policy often imply different things to different people, since our perceptions are colored by our varying ideologies, group affiliations, and economic interests. Plus, as you can see in Figure 1.1, these goals often conflict with one another. The most hotly debated trade-offs are the ones between equity and efficiency and between liberty and security.

However, before discussing the concepts of efficiency, equity, liberty, and security, we should also be mindful of the influence of another extremely important influencing factor: time. Time is important, since everything in life must be evaluated within a certain time period. For example, if I put some money away for my children's college education, I would seek investments where my benefits would peak around the time my children would be graduating from high school. In contrast, if I wanted short-term gains, I would use a different set of investment strategies. Similarly, the same person looking at a public investment at two different points in time may come to two radically different conclusions about its desirability. For instance, when the United States purchased Alaska from Russia for \$7.2 million in 1867, the move had severe critics, who saw Alaska as nothing more than a polar bear infested frozen wasteland. However, oil was soon discovered there, and during the Cold War period, Alaska's location was vital to U.S. efforts to monitor and check Soviet power; thus, the

FIGURE 1.1 Goals and Their Trade-offs



purchase of Alaska became accepted as a prudent use of public money.³³ In contrast, in 2003, when President George W. Bush announced the invasion of Iraq, some polls indicated that his decision was supported by an unprecedented nearly 90 percent of Americans. As U.S. forces became bogged down in a protracted war at a huge cost—both in lives and in money—public support evaporated quickly. With these caveats, let us briefly look at how our individual perceptions as well as group affiliation can color the way we look at public policies.

Efficiency

Of all the good things to which we aspire in public policy, efficiency seems to be the easiest to understand, as well as one of the most important. Efficiency is defined as the ratio between output and input—in other words, “are we getting our money’s worth?” This question of efficiency can be measured in two different ways. One way of determining efficiency is by an **absolute measure**: are we getting our money’s worth from a specific government expenditure? We can also ask if an expenditure is the most efficient use of money given all other needs. This is referred to as the **relative measure**, or the **opportunity cost**. Take for instance the example of the U.S. government’s decision to bail out the failing auto industry in Detroit. As Uncle Sam uses taxpayers’ money to shore up General Motors and Chrysler, many people ask whether these are wasted efforts, since much of their problem stems from the fact that U.S. consumers are increasingly purchasing foreign cars. This argument uses the absolute measure of efficiency. If we evaluate the efficiency of the bailout using a relative measure, we would examine the other ways in which the bailout money could be used; for example, that money could have been used to help out the states (which are in dire straits as a result of huge budget deficits), improve education initiatives, or repair crumbling infrastructure. The notion of opportunity cost in public policy debate is perhaps most pronounced in the allocation of money for NASA and the National Endowment for the Arts. Critics point out that while there are so many worthy causes and policy problems on Earth, it would be inefficient to spend money on exploration of the universe or promotion of the arts. These criticisms stem from the agencies’ inability to place monetary value on or produce physical evidence of their outputs; thus, it makes it difficult to weigh these outputs against expenditures.

The concept of efficiency is deeply ingrained in our everyday thinking. Yet when we look at it closely, we realize how it can look different from different perspectives, particularly when it comes to what we do in the name of the general public. Some of us assume that government is by nature inefficient, while others see it as a necessary instrument to correct the shortcomings of the market. The chasm between these two camps defines the difference in core beliefs between conservatives and liberals, Republicans and Democrats in the United States. As a result, we engage in endless debates about issues such as “pork barrel spending” (wasteful government spending to satisfy the special interests of politicians) and government efforts to help those who are poor and powerless (or even to act quickly to address a pressing need).³⁴ Given the sheer number of alternatives and divisions within the nation along the ideological divide, politicians communicate with their respective political bases with “glaring examples” of government waste of public money.



A CASE IN POINT

What Do You Mean by Government Efficiency, Equity, and Liberty?

In 2002 I was lecturing at prestigious Peking University to a group of students and government bureaucrats on the issue of efficiency in public policy. Before I started talking, I asked the audience to raise their hands if they thought the U.S. government was very efficient in implementing the goals of approved public projects. Most hands went up. I told them that in my hometown of San Diego, there is a significant traffic bottleneck. The two major highways run north and south, and there was a critical need to build a connecting road. The road, I told the audience, was placed in the regional master plan in 1958, when in China Mao Zedong's disastrous plan of social engineering, the Great Leap Forward, was beginning to take shape. Yet nearly forty-five years later, the road had not been constructed.³⁵ A stretch of about six miles had been built, but a series of lawsuits by environmental groups had stopped the construction of a one-mile section in the middle. In the meantime, the city and its population had expanded rapidly, and not only was the connecting road needed for ease of transportation, its absence also posed a serious security threat to citizens in one of the most fire- and earthquake-prone regions of the country. Without the connecting road, any quick evacuation plan would have been in serious jeopardy.

This example was particularly puzzling to the Chinese, who were looking to me, the U.S. expert, to tell them how to increase government efficiency. In a nondemocratic nation, government projects, from highway construction to building colossal dams for hydroelectric projects, are implemented immediately. Visitors to China are amazed at the breakneck speed with which buildings go up in its glittering, modern cities. And thanks to persistent double-digit income growth, the massive country, to its credit, has been able to cut its poverty rate in half over the past three decades. To many, therefore, China epitomizes government efficiency. Others, however, complain that there are hidden human costs in terms of lost liberty. Its critics argue that China's economic progress is possible because the Communist Party attempts to remain in power by silencing all opposition.³⁶

Discussion Points

1. Given this contrast, how do you define government efficiency? Do you think the Chinese government is more efficient than the U.S. government?
2. How does government strike a balance between efficiency and equity?

For instance, President Reagan, a conservative, evoked the image of a "welfare queen" in order to rally support during his 1976 campaign. According to Reagan, the welfare queen "has 80 names, 30 addresses, 12 Social Security cards and is collecting veterans benefits on four nonexistent deceased husbands. And she's collecting Social Security on her cards. She's got Medicaid, getting food stamps and she is collecting welfare under each of her names."³⁷ In contrast, during the presidency of George W. Bush, nothing

rallied his opponents like the picture of oil executives or the no-bid contracts granted to Vice President Dick Cheney's former associates at Halliburton to support the war effort in Iraq. These examples, to the minds of the faithful, served as perfect examples of government waste and inefficiency.³⁸ As you also can tell from these examples, the question of "objectivity" is a slippery concept.³⁹ Apart from outright corruption and mismanagement, the perception of efficiency often arises out of a group's ideological position.

Another important aspect of government inefficiency, and one that is often overlooked in the literature, is the ability of a democratic government to move quickly. This has been a source of heated debate among those who study the process of economic development and democratization in the countries of the developing world. As we contemplate the question posed in my example of highway construction in the United States (see "A Case in Point"), it shows that democracies are, by their very design, inefficient. This "planned inefficiency" is also known as the system of "checks and balances." Because no one has the supreme authority to push a nation to accept any single vision, democracies are often seen as vacillating, chaotic, inefficient, and mired in bureaucratic red tape. This debate has taken on a much bigger global context and rests at the core of a comparison between the processes of development in authoritarian China and democratic India.⁴⁰

Equity

There are few, if any, on Earth who oppose equity as a policy goal. The equity of every human is enshrined as a "self-evident truth" in the U.S. Declaration of Independence. But what did equity really mean to the Founders? In their nascent nation, women, slaves, and white men who did not own property were not allowed to vote. Although significant progress has been made in gender and racial equity, the meaning of equity remains vague today.

Suppose you are engaged in developing admissions standards for a very competitive state university. Since this is a state university and the bulk of its funding comes from the taxpayers of the state, should you make an immediate distinction between in-state and out-of-state students? This is a matter of equity based on *membership*. If I am a citizen of the state, whether I pay taxes or not—even if I am too poor to pay taxes—you may think that my candidacy to attend the school should be judged by different admissions and tuition standards.

The question of equity is also closely intertwined with the notion of *eligibility*: who is eligible to receive public goods? You could base eligibility on merit using a certain academic standard, which may include high school grade point average, the quality of a written essay, or standardized test scores. So, have you defined equity through defining merit? Not quite. In your effort to make admissions standards equitable, you will have to determine the relative weights you will place on the three areas of consideration. If you place too much emphasis on grade point average, you will be reminded that not all high schools have similar standards of grading and competition. A very high grade from a school with a poor performance record may not measure up against those from high-performing schools. Should you therefore take into account the relative standing of the schools? What about the essay? While we might think it surely reflects a student's ability to think critically and articulate properly, unfortunately some students get help from their parents, friends, or even paid "essay consultants." So, perhaps we should place the

most weight on SAT scores? Alas, even setting aside lingering doubts about the ability of SAT scores to predict a student's performance in college, standardized tests contain a number of systematic biases. Race, gender, ethnicity, religious affiliation, first-generation language skills, as well as parents' educational levels all play a role in determining the average test score of a candidate. You may take into account not only the current disadvantages, but also the lingering effects of past discrimination.

Finally, in your admission criteria you may consider applicants whose parents or siblings have attended or are currently attending the school. These "legacy" applicants receive special consideration, even when they don't meet the strict measures of merit. Since most of these students come from particularly successful parents who may have given a lot of money to the university, it may be considered not only a good business practice but also fair to admit them, even if their academic achievement doesn't measure up to that of other incoming students.

As you can see, each of these special considerations is bound to raise strong opinions on all sides. As a result, even a seemingly straightforward application of "equity" runs into all sorts of controversies.

The Equity-Efficiency Trade-off

During his campaign for president, Barack Obama vowed that he would give a tax break to 95 percent of American wage earners but would increase taxes on the top 5 percent, who earn more than \$250,000. The Republicans saw elements of "class warfare" in his position and bitterly complained that such a move amounted to a tax on "success." Their debate rests at the core of the trade-off between efficiency and equity.

Those on the right point out that the capitalist system is based on individual initiative and entrepreneurship in a free market. If I know that I am going to have to share my profits with the government, where is my incentive to take risks, work hard, and do better than everybody else? If we stifle such initiatives don't we strike at the heart of efficiency? If, by trying to redistribute income, we slow the pace of economic growth, then the equity in the distribution of income comes at the expense of market efficiency.

The left responds by saying that over the past twenty-five years, the income distribution in the United States has become exceedingly skewed in favor of the rich. Such a concentration of income not only gives economic power to a small group of oligarchs, it also gives them an undue share of political power. Furthermore, they argue that the money people can keep after taxes is not their only source of motivation. Our work is often the source of our satisfaction. We also have other needs, including the need to belong and to have a secure future not only for ourselves but also for our children. Thus when people work hard, they also seek these other nonmonetary ends, which may not hamper their desire to work hard even if they face higher taxes.

In contrast to those who see an inevitable trade-off between equity and efficiency, others have argued that when we look at economic growth over the long term, the apparent zero-sum game between equity and efficiency seems to disappear. Nations with a large middle class are inherently more stable than those with extremes of income.⁴¹ They further point out that Japan as well as many European countries enjoy high rates of economic growth despite having much higher tax rates (and more social programs) than the United States.

Liberty

"Give me liberty or give me death." In the face of an impending attack by British forces in 1775, Patrick Henry made this impassioned declaration and secured his firm place as one of the original American patriots. *Freedom* is one of the most overused terms in American political rhetoric, and the Statue of Liberty stands as the most recognized symbol of the entire nation. Yet our liberty is not absolute. Although we have the right to free speech, we cannot yell "Fire!" in a movie theater or smoke in many public places. The nineteenth-century English philosopher John Stuart Mill argued in his famous essay "On Liberty" that a man should be allowed to do whatever he wants to do—unless it harms others in the process. That is, the state must intervene and our liberty must be curtailed if in our pursuit of happiness we imperil others' well-being.

On the face of it, the notion of preventing someone from doing harm to others as a rationale for state interference seems fairly obvious. However, this harm is not confined to the physical integrity of a person or property. Think about the case of same-sex marriage or polygamy. Several states have passed referenda declaring marriage as a union solely between a single woman and a single man. These results are likely to work through the court system for several years as arguments are made for their unconstitutionality. If such a prohibition on same-sex marriage is upheld by the Supreme Court, it will in effect state that marriage as an institution will be harmed if we allow homosexuals to get married. Although there is no move to legalize multiple marriages—as are allowed in a number of Muslim countries—their prohibition makes the same statement about their harm.

How about harm to one's own self? It is a crime to try to commit suicide, since it is declared that the state has an abiding interest in our being, and thus taking one's own life would also harm the state's interest. What about doctor-assisted suicide in the case of terminally ill patients? Once again, you can see that no easy answer is available to address these tortuous questions.

Let us take another example. Should we legalize marijuana? *Time* magazine columnist Joe Klein has argued, in a somewhat lighthearted way, that with plenty of legal and illegal mind-altering drugs around, we should seriously consider legalizing marijuana.⁴² He points out that any harm to individuals may be overcome by the economic benefits of commercializing the drug, since it is estimated to be the largest cash crop in California, with an annual commercial value of \$14 billion. A 10 percent tax on marijuana could bring \$1.4 billion to the cash-strapped state. It also could create a large number of legal (and that means taxpaying) jobs in the agricultural and manufacturing sectors. And, more important, it could save the expenses of prosecuting and imprisoning a large number of drug offenders. On the other hand, the liberty of legally smoking marijuana and generating tax revenue must be weighed against a number of outcomes that can adversely affect society's welfare. For instance, the legalization and commercialization of marijuana would surely increase the number of users and enhance the legally produced drug's potency; neither of these consequences would be beneficial for society's health. Especially considering the global effort to reduce cigarette smoking, the introduction of another smoking agent makes little sense.

When discussing liberty and state sanction, we cannot overlook another extremely difficult question: if the authorities overstep their bounds and curtail our liberty, do we have a moral obligation to oppose tyranny? Mill was very aware of this issue and wrote

about *social tyranny* and the *tyranny of the majority*. Social tyranny takes place when the ruler(s) of a nation becomes despotic and makes decisions that harm the people, who have no way of escaping the harm. Mill also acknowledged that sometimes the majority of the people make decisions that are harmful to a small minority. In this “tyranny of the majority” scenario, the citizens have the right and even a moral obligation to oppose these policies. For instance, the Jim Crow legislation in the South, which discriminated against and oppressed African Americans, was based on consent of the majority of those whites who were living in southern states. Thus, although the laws might have had legal justification, based on every accepted standard of human rights, the laws’ utter injustice cried out for opposition.

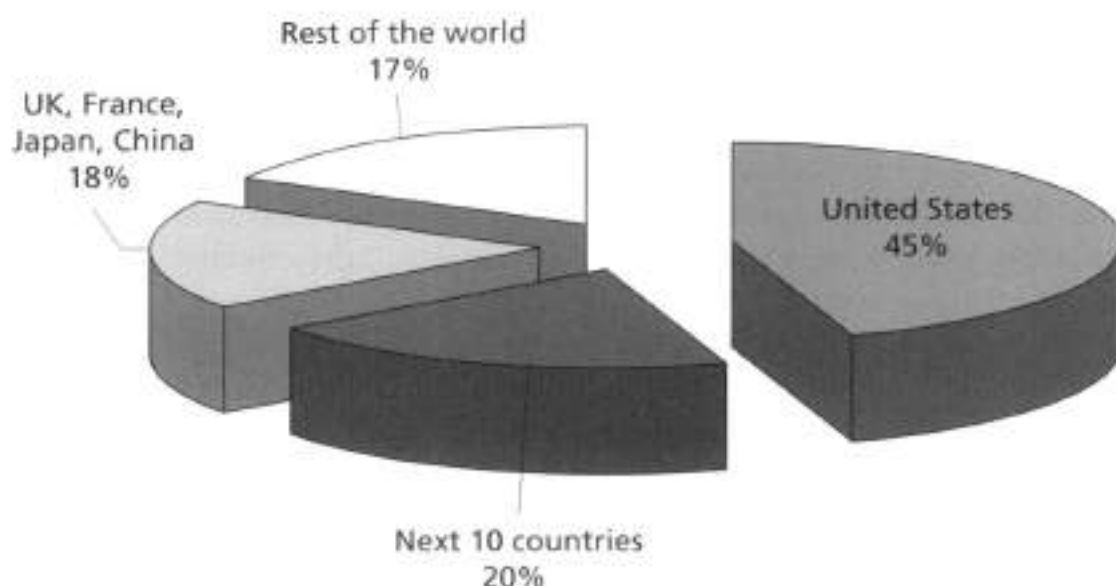
In sum, although Mill put his arguments in philosophical terms, in the real world, we can translate his theory into practice by emphasizing that, as in medicine, the first principle of public policy is to do no harm. We will have a chance to discuss this at length later in the book.

Security

Ever since humans began forming organized societies, kings sought to protect their subjects from external and internal threats. Security remains an important goal for governments. Even in the most libertarian interpretation of Adam Smith’s “invisible hand” and the primacy of market forces, the role of government in providing a secure environment is universally recognized. To do otherwise is to invite anarchy.

Security, however, is a very broad notion, encompassing physical security (which justifies military and police actions), as well as protection from natural disasters or ecological collapse. In recent years two images have clearly marked our collective memory with a strong sense of insecurity: the two hijacked planes that crashed through the twin towers of New York’s World Trade Center and the sight of New Orleans residents stranded on housetops and victims’ bodies floating in flooded streets after Hurricane Katrina. As a result, there are endless debates regarding how to allocate resources for the military and police. In 2008, the Stockholm International Peace Research Institute, a respected think tank, estimated that the total global military expenditure amounted to \$1.339 trillion, which is about 2.5 percent of the world’s gross domestic product, or \$202 for each person in the world.⁴³ The top fifteen countries with the highest spending account for 83 percent of the total expenditure, and the United States is responsible for 45 percent of the world total, distantly followed by the United Kingdom, China, France, and Japan, each with 4 to 5 percent of the world share (see Figure 1.2). Therefore, there is an ever-present debate in the United States about whether the government spends too much of its resources on the military. Those who stress security needs point out the necessity of maintaining a strong defense in an increasingly volatile world. In contrast, others point to the alternative uses of this vast sum of money—from the alleviation of poverty to education and health care.

External threats, particularly those posed by terrorism, traditionally have been the domain of the federal government. However, after the 9/11 attacks, the city of New York took a different tack and established its own intelligence unit.⁴⁴ Established by Raymond Kelly, the pragmatic commissioner of the New York Police Department (NYPD), and David Cohen, a former analyst at the Central Intelligence Agency, the NYPD’s counter-terrorism division, which stations more than 600 police officers and operatives stateside and around the world, is credited by some as the most effective force in the nation.

FIGURE 1.2 World Military Spending

Source: Stålenheim, Perdomo, and Sköns, "Chapter 5. Military Expenditure," *SIPRI Yearbook 2008*.

The demands of internal security in the United States also are overwhelming. By compiling data from the U.S. Department of Justice, the Pew Center on the States estimates that the total U.S. correctional population—those in jail, in prison, on probation, or on parole—stood at 7.3 million in 2007, or one in every thirty-one adults. Astoundingly, the United States has 5 percent of the world's population, but 25 percent of the world's prison inmates. This huge correctional population is also highly skewed in its composition; black adults are 4 times as likely as whites and nearly 2.5 times as likely as Hispanics to be in the prison system. Also, while the number of female offenders continues to grow, the number of imprisoned men (of all races) is 5 times that of women. This enormous prison population has created overcrowded facilities that have put many states' budgets in dire straits. This is particularly true in California, where the effects of "three strikes and you're out" legislation, in which three felony convictions result in exceedingly long mandatory prison sentences, have contributed greatly to the problem. Some have argued that overcrowding leads to inhumane treatment of prisoners, which ultimately contributes to their radicalization and thus increases threats to safety and even national security.⁴⁵

The Liberty-Security Trade-off

First and foremost, we desire security from our governments; this desire comes under sharper focus in times of crisis. After the 9/11 attacks, the federal government was forced into action, and many of the actions the George W. Bush administration took placed extraordinary power in the hands of the president and the executive branch. President Bush obtained the power to declare U.S. citizens "enemy combatants" and hold them indefinitely without trial, and the Department of Justice and the attorney general argued for the abrogation of the Geneva Convention, which protected detainees at Guantanamo Bay against torture.⁴⁶ Wiretaps were also authorized without any judicial oversight. These extraordinary steps were justified in the name of providing security to the

American people against future attacks. When the nation felt insecure, the general public accepted these assaults against individual liberties as necessary steps. As the threat of a further attack seemed to fade and the wars in Iraq and Afghanistan turned into quagmires, however, resistance to these security measures built up among the public.

When we define security in the traditional sense of managing internal and external threats, we can clearly see a trade-off with liberty. Recently, however, a group of notable scholars have been promoting the concept of **human security**. In 1994, Pakistani economist Mahbub ul Haq crystallized the concept of human security in a United Nations Development Programme report, defining the term according to seven factors:

1. *Economic security*—A minimum income through meaningful employment or from government assistance.
2. *Food security*—The assurance of enough food for everyone.
3. *Health security*—Protection from diseases and unhealthy living conditions, including access to clean drinking water.
4. *Environmental security*—Protection from natural or manmade disasters.
5. *Personal security*—Protection to individuals who are threatened by violence from ethnic and religious conflicts or criminal groups.
6. *Community security*—Protection to groups that are threatened by ethnic, religious, cultural, or political conflicts.
7. *Political security*—The ability to express political beliefs without fear of repression from governments or other groups.

In a 2003 United Nations Commission on Human Security report, the economist Amartya Sen argued that the concept of security is meaningless without freedom and people's ability to have economic security, health security, and freedom.⁴⁷ To Sen, security, economic development, and freedom constitute parts of a larger whole.⁴⁸ In this broader definition of security, it is not a zero-sum game; there is indeed no trade-off between security and liberty.

As we have seen throughout this chapter, public policy analysis involves intricate balancing on multiple levels. Governments, for one, must balance notions of liberty and security, efficiency and equity, in order to maintain a functioning society. As political beings, all of us also perform a balancing act, weighing the self-interested impulses of our inner *homo economicus* against our subjective and collective perceptions. Within this complex puzzle of rationality, human behavior, and government action, public policy analysts work to sort out the pieces and find the best ways in which they fit together.

KEY WORDS

Absolute measure (p. 16)	Exhaustibility (p. 6)
Alternatives (p. 4)	Falsifiability (p. 5)
Club good (or common pooled resource) (p. 7)	Framing (p. 13)
Collective action problem (p. 8)	Free-rider problem (p. 8)
Credible commitment problem (p. 8)	Human security (p. 23)
Efficiency (p. 15)	Invisible hand (p. 3)
Equity (p. 15)	Liberty (p. 15)
Excludability (p. 6)	Model (p. 5)

Moral hazard (p. 8)
 Opportunity cost (p. 16)
 Policy cycle (p. 3)
 Political entrepreneurs (p. 13)
 Politics (p. 2)
 Private good (p. 6)
 Prospect theory (p. 14)
 Public good (p. 6)

Pure public goods (p. 7)
 Rationality (p. 2)
 Relative measure (p. 16)
 Rivalrous (p. 6)
 Security (p. 15)
 Shirk (p. 8)
 Toll good (p. 7)
 Tragedy of the commons (p. 9)

EXERCISES

1. Why is it difficult to define human rationality? Why is understanding rationality important for the discussion of public policy?
2. What are the major ethical trade-offs in the realm of public policy? Consider government responses to recent catastrophic events, such as the 9/11 attacks, Hurricane Katrina, or the economic meltdown of 2008. At times of stress, how do we, as a society, deal with the contradictions between equity and efficiency or liberty and security?
3. What is human security? How is it defined, and how does its acceptance reduce conflict with equity and efficiency?
4. A report on poverty in the United States prepared by the U.S. Government Accountability Office (GAO) appears in Appendix A. The report explains how poverty creates its own vicious cycle. Read the report and discuss in a group what the role of government in ending this cycle of poverty should be. Explain how individuals with different political ideologies might interpret the results differently.
5. How does the concept of human security relate to the GAO's report on poverty?

Notes

1. Adam Smith, *An Inquiry into the Nature and the Causes of the Wealth of Nations* (New York: Modern Library, Random House, 1937), 461. (Originally published in 1776.)
2. Walter Isaacson, "How to Raise the Standard in America's Schools," *Time*, April 27, 2009.
3. It is interesting to note that while Adam Smith's assumption of self-utility-maximizing human being, proposed in *The Wealth of Nations*, gained widespread acceptance and became the foundation of modern economics, he had argued for a much more nuanced view of human nature in his lesser known earlier work *The Theory of Moral Sentiments*.
4. French philosopher Michel Foucault calls the assumption of me-centric individualism "positive consciousness of knowledge," which he defines as "a level that eludes the consciousness of the scientists yet is a part of scientific discourse." Michel Foucault, *The Order of Things: An Archeology of the Human Sciences* (New York: Pantheon Books, 1970), xi.
5. George Lakoff, *The Political Mind: Why You Can't Understand 21st-Century American Politics with an 18th-Century Brain* (New York: Viking, 2008), 3.
6. Amartya K. Sen, "Rational Fools: A Critique of the Behavioral Foundation of Economic Theory," *Philosophy and Public Affairs* 6 (1977): 317–344 (italics in the original).
7. See Elinor Ostrom, "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997," *American Political Science Review* 92 (March 1998): 1–22.

8. Bengt Holmstrom, "Moral Hazard in Teams," *The Bell Journal of Economics* 13 (Autumn 1982): 324–340.
9. Armen Alchian and Harold Demsetz, "Production, Information Costs and Economic Organizations," *American Economic Review* 62 (1972): 777–792.
10. Gary Miller, "Credible Commitment and Efficiency in the Design of Public Agencies," *Journal of Public Administration Research and Theory* 10 (2000): 289–328.
11. Garrett Hardin, "The Tragedy of Commons," *Science* 162 (December 1968): 1243–1248.
12. Mancur Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, Mass.: Harvard University Press, 1965), 161–162.
13. For an alternative formulation of human nature in social sciences, see Dipak K. Gupta, *Understanding Terrorism and Political Violence: The Life Cycle of Birth, Growth, Transformation, and Demise* (London: Routledge, 2008). Also see Dipak K. Gupta, *Path to Collective Madness: A Study in Social Order and Political Pathology* (New York: Praeger, 2001).
14. David Brooks, "Peering into Our Economic Future," *New York Times*, February 13, 2009.
15. Antonio R. Damasio, *Descartes' Error: Emotion, Reason, and the Human Brain* (New York: Avon, 1994). Also see Daniel Goleman, *Emotional Intelligence* (New York: Bantam, 2006).
16. Drew Westen, *The Political Brain: The Role of Emotion in Deciding the Fate of the Nation* (New York: Public Affairs, 2007), x–xv.
17. *Ibid.*, xiii–xv.
18. *Ibid.*, xiv (italics in the original).
19. It may interest you to note that some cognitive psychologists argue that we are predisposed to our ideologies. One study concludes that your answers to two completely unrelated questions can predict your ideological orientation. The first question is, How willing would you be to slap your father (with his permission) as a part a comedy skit? The second is, How disgusted would you be to touch a faucet in a public restroom? The study's authors argue that your answers to similar questions can place you on a scale between extreme conservatism and extreme liberalism. They contend that the fundamental difference between the two rests on one's willingness to accept authority and the feeling of disgust one feels in various life situations. Conservatives have greater respect for authority and feel greater disgust than liberals. Visit www.yourmorals.org to take the test and find out your own value position.
20. Pew Research Center for the People & the Press, "Partisan Gap in Obama Job Approval Widest in Modern Era," April 2, 2009, <http://pewresearch.org/pubs/1178/polarized-partisan-gap-in-obama-approval-historic>.
21. See Robin Dunbar, Louise Barrett, and John Lycett, *Evolutionary Psychology* (Oxford, United Kingdom: Oneworld, 2005).
22. For an amusing yet frightening discussion of English soccer fans, see Bill Buford, *Among the Thugs* (New York: Norton, 1992).
23. Benedict Anderson, *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (New York: Verso, 2003).
24. For an explanation of our altruistic behavior, see Kristen Renwick Monroe, *The Heart of Altruism: Perceptions of a Common Humanity* (Princeton, N.J.: Princeton University Press, 1996).
25. By expanding Joseph Schumpeter's concept of an economic entrepreneur from *The Theory of Economic Development*, the noted economist William J. Baumol made a useful distinction between "constructive" and "destructive" entrepreneurs, whose behaviors are guided by the reward system within a society. See William J. Baumol, "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy* 98 (October 1990): 893–921.
26. For one of the best explanations of "irrationality" due to framing, see James N. Druckman, "The Implications of Framing Effects for Citizen Competence," *Political Behavior* 23 (September 2001): 225–256. Also see James N. Druckman, "Political Preference Formation:

- Competition, Deliberation, and the (Ir)relevance of Framing Effects," *American Political Science Review* 98 (2004): 671–686.
27. See Samuel Kernell and Gary C. Jacobson, *The Logic of American Politics* (Washington, D.C.: CQ Press, 2000).
 28. See David Ignatius, "Bernanke, or New Tiger atop the Fed?" *Washington Post*, May 28, 2009.
 29. See Rose McDermott, James H. Fowler, and Oleg Smirnov, "On the Evolutionary Origin of Prospect Theory Preferences," *Journal of Politics* 70 (2008): 335–350.
 30. Daniel Kahneman and Amos Tversky, "Choice, Values, and Frames," *American Psychologist* 39 (1984): 341–350.
 31. Jonathan Haidt, *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom* (New York: Basic Books, 2006), 29.
 32. For a delightful discussion of our lemming-like behavior, see Robert B. Cialdini, *Influence: How and Why People Agree to Things* (New York: Morrow, 1984).
 33. Ronald J. Jensen, *The Alaska Purchase and Russian-American Relations* (Seattle: University of Washington Press, 1975).
 34. For an interesting discussion of legislators' motivations behind "bringing home the bacon" ("pork barrel" to others), see Philip Keefer and Stuti Khemani, "When Do Legislators Pass on Pork? The Role of Political Parties in Determining Legislator Effort," *American Political Science Review* 103 (2009): 99–112.
 35. For a history of Highway 56 in San Diego, see "Highway 56 History: 1958–Present," July 18, 2004, www.gbcnet.com/roads/hwy_56/56_history.html.
 36. See, for example, Elizabeth C. Economy and Adam Segal, "China's Olympic Nightmare: What the Games Mean for Beijing's Future," *Foreign Affairs*, July/August 2008.
 37. "'Welfare Queen' Becomes Issue in Reagan Campaign," *New York Times*, February 15, 1976. For a contrary view, see Susan J. Douglas and Meredith W. Michaels, *The Mommy Myth: The Idealization of Motherhood and How It Has Undermined All Women* (New York: Free Press, 2005).
 38. Despite widespread press, there were many who saw the case against Halliburton differently. See, for instance, Byron York, "Halliburton: The Bush/Iraq Scandal That Wasn't," *National Review*, July 14, 2003.
 39. For an excellent discussion of the subjective nature of government's participation in the market, see Deborah Stone, *Policy Paradox: The Art of Political Decision Making* (New York: W. W. Norton, 2002).
 40. See, for example, Daniel Deudney and G. John Ikenberry, "The Myth of the Autocratic Revival: Why Liberal Democracy Will Prevail," *Foreign Affairs*, January/February 2009.
 41. See Dipak K. Gupta, *The Economics of Political Violence: The Effect of Political Instability on Economic Growth* (New York: Praeger, 1990).
 42. Joe Klein, "It's High Time: Why Legalizing Marijuana Makes Sense," *Time*, April 13, 2009.
 43. Petter Stålenheim, Catalina Perdomo, and Elisabeth Sköns, "Chapter 5. Military Expenditure," *SIPRI Yearbook 2008*, <http://yearbook2008.sipri.org/05>.
 44. Christopher Dickey, *Securing the City: Inside America's Best Counterterrorism Force—The NYPD* (New York: Simon & Schuster, 2009).
 45. See Mark S. Hamm, *Terrorism as Crime: From Oklahoma City to Al-Qaeda and Beyond* (New York: New York University Press, 2007).
 46. For an excellent discussion, see Jane Mayer, *The Dark Side: The Inside Story of How the War on Terror Turned into a War on American Ideals* (New York: Doubleday, 2008).
 47. Commission on Human Security, *Human Security Now*, 2003, <http://www.humansecurity-chs.org/finalreport/English/FinalReport.pdf>.
 48. See Amartya Sen, *Development as Freedom* (New York: Knopf, 2009).

2

The Analysts: Their Role and Their Tools

The field of public policy is so vast that those who study it may be compared to a group of blind people who have been asked to describe an elephant: some people describe it by its trunk, some by its tusks, some by its large belly, some by its large, floppy ears or by its thin tail. According to the University of California's electronic library catalog, Melvyl, no fewer than 4,586 books with the words *public policy* in their titles have been published during the past ten years. This enormous number of publications clearly indicates that there are many ways of studying public policy.

While some students of public policy focus on the political, cultural, and judicial contexts in which policies are made, others study how policies are adopted. Economists concern themselves primarily with the effects of policies, while operations researchers and those who use quantitative techniques examine the management policies that promote efficiency in public organizations' delivery of services. Philosophers worry about the ethical aspects of government involvement in our everyday lives. Students of public administration examine the role of bureaucracy in the allocation of public goods. Political scientists, in contrast, typically are most interested in the role of interest groups in public policy.

If you define the study of public policy as advising those who govern, the field has a long history.¹ If, on the other hand, you define it as a systematic, objective, and institutionalized approach to improving the art of government, then policy science has a short past. Let us keep in mind both definitions as we discuss the evolution of the profession and the tools developed by policy analysts.

THE PRINCE AND THE PUNDITS: ANALYSTS, ACADEMICS, AND ADVOCATES

Throughout history, kings and princes depended on counsel from their ministers. These ministers were the first policy analysts, since they were the advisers to the ultimate authorities (and the actual decision makers) in the land. Today, in a large democracy, the role of policy analysts is much more diffuse. Although it is not always clear whom we should call a "policy analyst," this large group may be divided into *staff analysts*, *academics*, and *advocates*.

Policy analysts are typically employed by governments to help executive decision makers at all levels of government—federal, state, and local—on issues ranging from budgeting to law enforcement to health and human services (see "A Case in Point"). They also may be employed by the legislative and the judiciary branches of the government or by multinational agencies, such as the World Bank, the International Monetary Fund, the United Nations Children's Fund, the World Meteorological Organization, and the International Labour Organization.² Policy analysts are experts in their fields, and often become repositories of information. They collect reports written by other agencies, articles in professional journals, and scholarly books in their areas of expertise, and become the "go-to persons" for the decision makers in all branches of the government.

Apart from those who are directly employed by governments, many analysts work as hired consultants. They can be academics or employees of private consulting firms or nonprofit organizations. These experts are typically hired to work on specific projects for which government agencies do not have "in-house" capability. For instance, suppose your city's administration has noted that fatal car accidents along a certain corridor have increased significantly, and they need a plan to reduce this number. If there is no one in the transportation department to conduct the study, the department may issue a request for proposal and hire an individual contractor or a firm to do the job. Applicants for the job might include experts in the field, former transportation officials, consulting firms, private think tanks, or professors teaching transportation planning or traffic engineering at your university. Government-funded institutes (as opposed to those directly run by the government), such as the National Science Foundation, the National Institutes of Health (NIH), and the United States Institute of Peace, may also hire policy analysts to evaluate specific programs. For example, the NIH may fund a program to evaluate the relative success of various antismoking campaigns among teens.

Finally, the world of policy analysts is also composed of nongovernmental organizations that approach policy problems from a strict advocacy position. As opposed to the analysts working for government agencies or institutes, advocacy analysts are part of a large network of interest groups. Their clients are specific interest or ideological groups, and their effort is ultimately directed toward the benefit of their clients. For instance, the Tobacco Institute, which was founded by the powerful tobacco industry in 1958, for decades produced or funded studies that fraudulently disclaimed any link between smoking (and exposure to secondhand smoke) and lung cancer, hypertension, and other forms of cardiovascular disease.³ Similarly, in the political arena, some research institutes have clear ideological orientations. For example, the Cato Institute, the American Enterprise Institute, and even the Stanford University-affiliated Hoover Institution are seen as "right wing," while the Brookings Institution is regarded as "liberal." As a result, analyses conducted by these think tanks often carry signs of an ideological slant.



A CASE IN POINT

The Numb and Number: The Exploding Domain of Public Policy Analysts

Since the government allocates money, which is spent—at least partially—according to what policy specialists recommend, the domain of policy analysts is enlarging both in relative and in absolute terms. The amount of government spending is so huge that the numbers can truly numb our minds. As I write this book in 2009, the United States is in the grips of a deep recession. As the number of foreclosed homes rises and many people find themselves out of work, newly elected president Barack Obama has attempted to jump-start the economy with a direct infusion of public money. The question, however, is how much money?

We all know that a million dollars is a lot of money. Each year *Forbes* magazine lists the wealthiest people in the world, and their worth is now counted in billions. However, when it comes to the federal government, the operative word is *trillion*. Facing the daunting task of restarting the economic engine, President Obama proposed a budget of \$3.6 trillion. At this moment, the total federal debt—the amount Uncle Sam owes to those who have lent him the money—stands at an astounding \$10 trillion.

How large is the number 1 trillion? If I were to ask you what you were doing a million seconds ago, would you be able to remember? You should, because a million seconds equals approximately 11.5 days. Now, let me ask you what you were doing a trillion seconds ago. Well, if you cannot recall, there is a good reason for it. Since a trillion is a million millions, a trillion seconds translates into 11.5 million days—or about 32,000 years.

These trillions of dollars in government expenditures, however, do not fall from the sky. Like any borrower, the government must repay its debts. Every time the government undertakes a new program, it creates “winners” and “losers”: those who benefit from it and those who pay for it. Given the enormous size of the government’s expenditures, the question therefore is how to make sure the money is spent efficiently and with due consideration to justice and equity. These questions constitute the domain of the public policy analyst. As the size of the government increases, so do the responsibilities of those who must make sure that its funds are being spent wisely.

Discussion Points

1. What do you think is the appropriate role of government in a free market society?
2. Does the overwhelming amount of public spending make the job of policy analysts more important? Explain.

The ranks of policy analysts also include journalists and columnists, who often take the side of a particular political party or a specific ideology. You could take some of the best-known columnists and, based on their articles or media appearances, place them along the ideological spectrum.

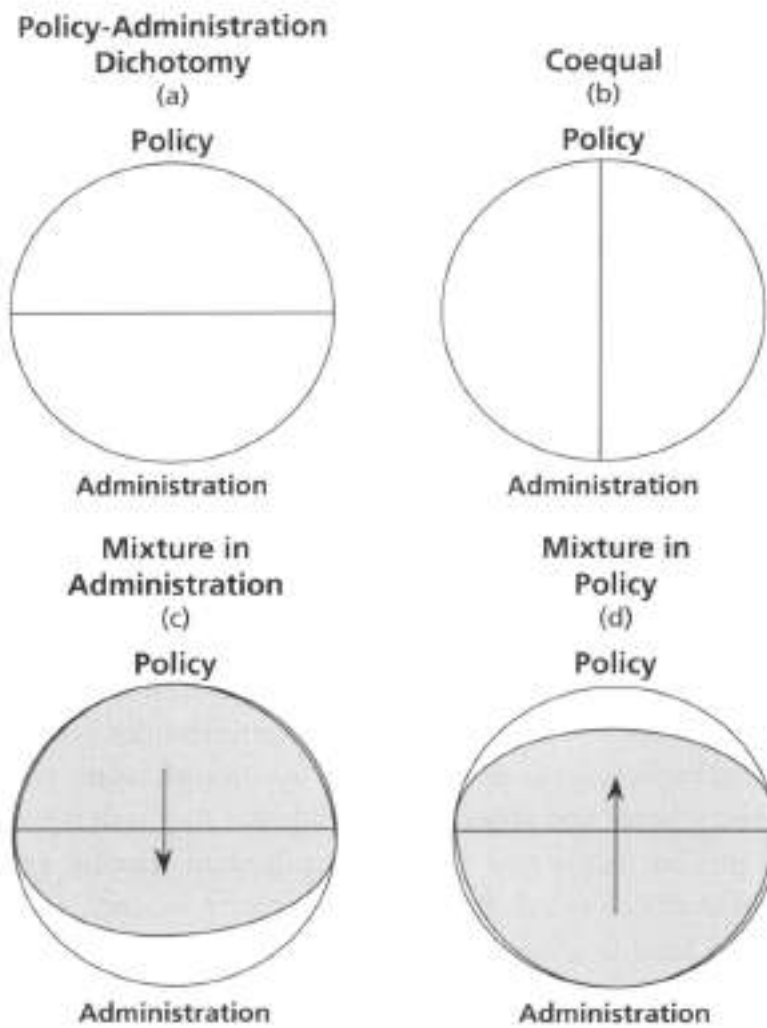
Policy analysts generally serve specific client groups. The idea of a completely impartial, rational, scientific analysis is often a figment of the imagination in an otherwise subjective world. Therefore, it is hardly surprising that within the democratic system, the role of unelected bureaucrats in general, and staff policy analysts in particular, will be of great interest. Some have accused these bureaucrats of having too much influence, while others have placed greater faith in these trained professionals than in the elected “political animals.” Let us therefore further examine the complicated role of policy analysts in a democratic system.

The Politicos and In-house Policy Analysts: An Uneasy Cohabitation

The Framers of the Constitution minimized the role of the government, particularly at the federal level. However, as the new nation took on greater responsibilities on an ever-larger scale, government officials recognized the need for true professionals to run the government and make its policies. In the early years, the administration and politics were inseparable, and elected politicians hired their own supporters to fill administrative positions. However, as the government grew more complex, untrained administrators drew criticism for mismanagement and corruption, which prompted a public call for change.

The first serious study of public administration and the role of the bureaucracy began three years after the passage of the Pendleton Act (1883), which sought to reform the civil service by instituting merit-based selection procedures. The author of the study, a young political scientist named Woodrow Wilson, stressed the importance of a strict separation between elected officials and nonelected bureaucrats.⁴ His model advocated a perfect dichotomy between the two groups: elected officials defined policy objectives while public agencies and administrators performed purely technical and professional functions. Policymakers did not interfere in administration, and administrators did not get involved in policymaking. The horizontal line bisecting the sphere in Figure 2.1 (a) shows this strict hierarchy: elected officials make public policies, and administrators remain fully accountable to their political bosses while translating adopted policies into action. This **policy-administration dichotomy model** primarily emphasizes democratic control over a burgeoning bureaucracy. Insulated from the corrupting influences of the political spoils system, in which elected officials ignore the merits of job applicants and offer administrative posts to their loyal supporters, administrators work toward achieving greater degrees of neutrality, professionalism, and efficiency. Introduced in the 1930s, this model withstood strong attacks from various academic quarters to become, until the mid-1960s, the dominant model in the field.

In 1968, during a tumultuous period in American history, a group of young academics founded the “new public administration,” a school of policy analysis. Like most young men and women of that time, they were deeply affected by the riots, the assassinations of prominent political leaders, and the ongoing war in Vietnam, which fostered in them a strong distrust of elected officials. These scholars distanced themselves from the policy-administration dichotomy model, which seemed to them to prescribe a machinelike devotion by administrators to the policies of elected officials.⁵ They also questioned the loyalty of administrators to their political bosses.⁶ They

FIGURE 2.1 Four Models of Policymaking

Source: Based on James H. Svara, "Dichotomy and Duality: Reconceptualizing the Relationship between Policy and Administration in Council-Manager Cities," in *Ideal and Practice in Council-Manager Government*, ed. H. George Fredrickson (Washington, D.C.: International City-County Management Association, 1995), 3–19.

advocated a nonhierarchical, coequal arrangement in which administrators were loyal not to the hierarchy but to their ethical principles—their belief in the equality of policymakers and administrators and their commitment to protecting the rights of the powerless.⁷ Radical in their approach, they urged administrators to become activists by developing a direct link with the people. The young academics admitted that such direct contact might sometimes conflict with the policies of elected representatives. Figure 2.1 (b) illustrates the **coequal model**: a vertical line divides the policy-administration sphere into two equal domains.

Another group of scholars went in the opposite direction in the 1960s. They developed models in which elected officials questioned the role of administrators in policymaking. The **mixture-in-administration model** (see Figure 2.1 (c)) observes that elected officials prevent nonelected administrators from gaining too much authority in a democracy by encroaching on the domains of administrators. At the national level,

elected officials encroach on administrative domains when different parties control the White House and the legislative branch. For example, the Democratic-controlled 110th Congress that came to power in 2006 called Bush administration officials to testify about their operations. As U.S. forces became bogged down in Iraq and the need to take quick and decisive action against terrorism was replaced with calmer reflections, the Bush administration was forced to defend its actions at home and abroad. The use of the USA PATRIOT Act of 2001 (Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism) and allegations of torture at the Guantanamo Bay detention camp became the preoccupation of numerous congressional oversight committees.

In contrast to the mixture-in-administration model, the **mixture-in-policy model** (see Figure 2.1 (d)) describes the increasing role of bureaucracy in public policy making.⁸ Three developments fostered this rapid encroachment by nonelected policy analysts and executives with specific expertise, particularly in local governments. First, as the nation grew more populous and urban, the demand for public services increased more than proportionately. New demands for public policies on crime control, education, health care, infrastructure, drug abuse, poverty, and environmental protection particularly challenged city governments, which typically were ill equipped to deal with such a wide variety of issues. Second, at the federal level, the Great Depression and Keynesian economics, which prescribed an involved government, changed people's attitudes. People started expecting the government to act on such issues as raising children's test scores in public schools and reducing gun violence. As a result, public administrators and elected officials faced new rules of engagement. Finally, rapid advances in technology—from stem cell research to invasive security procedures at airports, from genetically modified food to climate change and global warming—required that government regulate and promote activities that the Founders never could have imagined.

Exit, Voice, and Loyalty: What's an Analyst to Do?

From the earliest days of recorded antiquity, the relationship between princes and high priests has been like a marriage, with alternating periods of gleeful bliss and contentious acrimony. Although the princes have depended on the pundits to chalk out the proper course of action, history is clear about the fate of those counselors who gave their masters advice they did not want to hear. King Henry VIII, although respectful of his advisers Cardinal Thomas Wolsey⁹ and Sir Thomas More,¹⁰ did not let his admiration of these men stand in the way of virtually imprisoning Wolsey and beheading More. Policy analysts do not suffer the same fates in today's world, yet at times they may experience a deep ambivalence working within an organization (see "A Case in Point").

When conflicts based on value and ethical considerations become deep, analysts face three broad choices. They can voice dissenting opinions, quit their organizations, or become disloyal to them. Based on the seminal work of Albert Hirschman,¹¹ Weimer and Vining¹² presented these three choices with three intersecting circles that illustrate seven different courses of action (other than simply being quiet and going along with one's supervisor) (see Figure 2.2). When facing a conflict with your organization, your first option is to voice dissent and *protest*. In this option you voice your opinion



A CASE IN POINT

Ignored Advice and the Global War on Terror

As a seasoned veteran of the intelligence community who had advised three presidents, Richard Clarke had spent years following the activities of a Saudi millionaire named Osama bin Laden and the radical Islamic fundamentalist group that he established, al-Qaida. Clarke linked the group to multiple attacks against the United States, including the 1998 bombing of the U.S. embassies in the East African cities of Dar es Salaam, Tanzania, and Nairobi, Kenya. He was convinced that the same group was responsible for the attack on the warship U.S.S. *Cole* off the coast of Yemen in 2000. And by 2000, he strongly suspected that al-Qaida was about to launch a devastating attack on U.S. soil. In fact, he had persuaded the Clinton administration to carry out missile attacks on bin Laden's alleged hideout in Afghanistan and a suspected bomb factory in Sudan. Clarke maintains that President Clinton was aware of the gravity of the situation and became increasingly obsessed with stopping al-Qaida. With Clarke, the Clinton administration had developed workable plans to neutralize bin Laden and his group but was thwarted by political infighting and the sex scandal that led to Clinton's impeachment by the House.

When the Bush administration took charge, they were skeptical of Clarke's claim that a ragtag group was targeting the United States. Some of the new politicians in town, known as the neoconservatives, or neocons, came in with the notion that it was the Iraqi dictator Saddam Hussein who should be the primary target of U.S. counterterrorism policy, and they aimed to change the political landscape of the Middle East by tumbling Hussein's regime. As Clarke tried in vain to convince Vice President Dick Cheney, Defense Secretary Donald Rumsfeld, and others in the administration of the ominous threat, he was told to concentrate on Saddam Hussein rather than wasting his time on bin Laden.

After al-Qaida's devastating attacks on the United States on September 11, 2001, the political powers in Washington kicked into full gear their plans to invade Iraq and take out Saddam Hussein. President Bush and Vice President Cheney disregarded the adversarial relationship between bin Laden and Hussein and made every effort to lump the two enemies together. As a frustrated Clarke watched, the United States was soon embroiled in a two-front war in Afghanistan and Iraq.

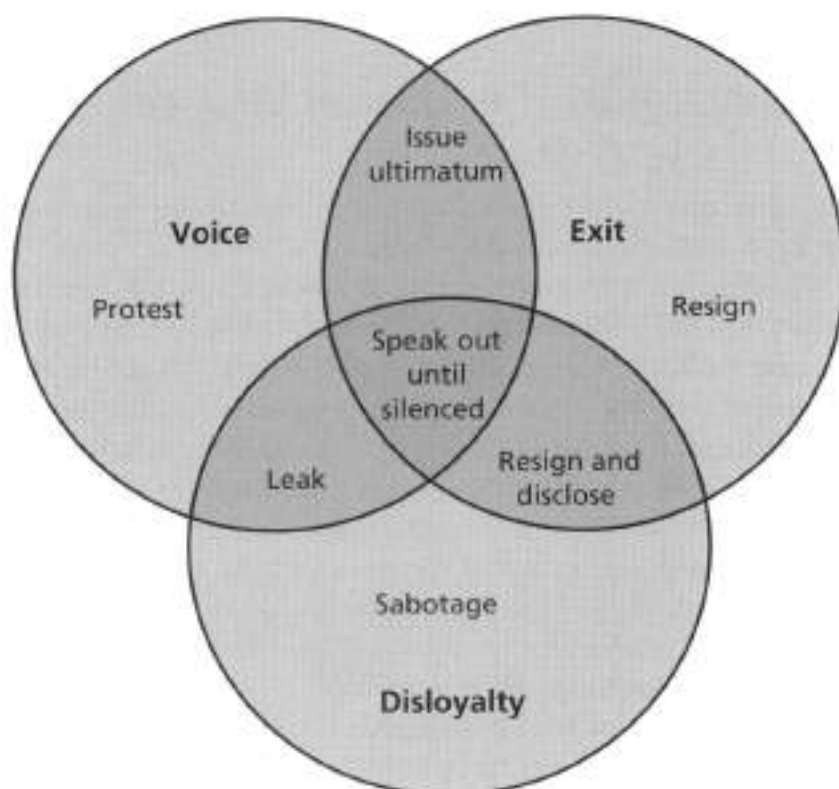
Discussion Points

1. When a politician ignores an analyst's recommendation, who is responsible for the consequences?
2. What would you have done if you were in Richard Clarke's position?

internally and do nothing to undermine the organization by going public with your grievances. If you feel that your position is being sufficiently undermined, you may be ready to exit and submit an *ultimatum*; unless the organization alters its position, you will leave. If you don't want to jeopardize your position within the organization, you

FIGURE 2.2

Alternative Responses to Value Conflict



Source: From David L. Weimer and Aidan R. Vining, *Policy Analysis: Concepts and Practice*, 3rd edition, Upper Saddle River, N. J.: Prentice-Hall, 1999. © 1999 by Prentice Hall Inc. Reprinted by permission by Pearson Education Inc.

may secretly *leak* information to outsiders. The risk, of course, is that your involvement may be discovered, and punitive actions may be taken against you. You may not want to remain loyal to the organization and thus may talk to the media or other outside departments. If you do, this would be regarded as a case of disloyalty. If you choose to be disloyal, perhaps the most honorable course of action is to *speak out* until you are silenced. People who do this are often called **whistleblowers**. A number of state and federal laws have been enacted to protect whistleblowers. If you are not willing to become a whistleblower, you may *resign* from your position and, once out of the organization, *disclose* its inner workings by writing a “tell-all” book. This has been the time-honored path of many in positions of power in Washington and elsewhere. Facing a moral dilemma, they carefully accumulate evidence, keep secret memos, and, after resigning as loyal followers, decide to publish memoirs for considerable sums of money. Finally, you may decide to *sabotage*. This typically happens when a disgruntled worker, finding no other option to stop an organization from moving forward in a way that the worker finds abhorrent, decides to work secretly against the organization’s goals. Most often these acts are hidden from the outside, since the organization and its supporters are too strong, or the laws do not protect those who might otherwise have come out in the open and opposed its policies. These saboteurs, when caught, can be severely reprimanded, and their stories often are fodder for Hollywood dramas (see “A Case in Point” for examples of these courses of action).

THE USE OF QUANTITATIVE TECHNIQUES IN PUBLIC POLICY ANALYSIS

In the un-rational world separated by group identities and ideological divisions, the search for scientific objectivity in public policy analysis has been a matter akin to finding the Holy Grail. From the beginning, the founders of policy analysis as a distinct field of inquiry—defined by Robert Lynd in the 1930s and 1940s and by Harold Lasswell in the 1940s and 1950s—sought to improve government efficiency by making better use of knowledge from the social sciences.¹³ Over the next half century, numerous scholars established the academic credentials of policy analysis, and today nearly all major universities either offer stand-alone academic degrees in public policy or offer the subject as an emphasis within an established program.¹⁴ As governmental organizations have striven to become more efficient by using mathematical models and quantitative techniques, the new fields of decision sciences and operations research have emerged.¹⁵ Together, they have come to be known as the **rational model** of decision making.

One of the primary drawbacks of mathematical reasoning is that it often gives the impression of objectivity and, as a result, raises expectations far beyond what it can possibly deliver. For example, on August 12, 1965, President Lyndon Johnson issued an executive order requiring every agency to institute what he called a “revolutionary” new system based on a systems approach—the Planning-Programming-Budgeting System (PPBS).¹⁶ Before the PPBS was implemented, the traditional federal budgeting process was incremental: each department’s budget request was considered individually in light of its allocation from the previous year. For example, if the Department of Defense submitted a budget request under the incremental system, policymakers focused on how much the department had received the previous year and asked whether the department could keep up with inflation and other relevant demands. Then, after complex political bargaining, legislators allocated the money. What was lost in this process, however, was the question of whether any of the programs were worth continuing. In contrast, the PPBS used a series of rigorous steps—such as defining the overall goal, finding alternative plans for achieving the goal, developing productivity indicators, and using a decision model such as cost-benefit analysis—to determine department budgets, ignoring how much funding an office had received in previous years. In this way, its proponents claimed that the PPBS eliminated “politics” from the budget process and allocated more resources to the most productive programs while curtailing or even eliminating those that were inefficient. In his statement introducing the PPBS, President Johnson promised that it would “improve our ability to control our programs and our budgets rather than having them control us.” This new system, bristling with the tools of the objective decision sciences, came to be known as the rational model of decision making. Pinning the term *rational* on one method made all other tools seem, at least implicitly, “irrational.” Two decades later, President Ronald Reagan issued an executive order directing that lawmakers analyze the costs and benefits of every federal regulation before recommending it for adoption.¹⁷

Despite the enthusiasm of President Johnson and many others in the policy field, the PPBS failed to meet expectations. Decision makers, doubting their ability to control the programs and budgets, instead were controlled by something they did not fully understand. As political scientist Aaron Wildavsky emphatically stated, “No one knows how to do program budgeting (PPBS).”¹⁸ Moreover, Wildavsky claimed that a “rational”



A CASE IN POINT

The Saboteur, The Whistleblower, and the Loyal Protester

They were some of the darkest days of American history, in early 1973, when a paranoid president used his power to subvert the democratic processes of the nation. As President Nixon faced widespread criticism for his increasingly futile war efforts in Vietnam, he and his advisers resorted to all sorts of illegal means—from unauthorized wiretapping and espionage to breaking into the Democratic Party's Washington office—to silence their opposition. Two young *Washington Post* reporters, Bob Woodward and Carl Bernstein, began investigating the administration's scandalous actions, and they were aided by a political informant whom they dubbed Deep Throat. Hiding beneath a broad-brimmed hat and behind the upturned collar of his trench coat, Deep Throat would meet the reporters in an empty parking garage or other secret locations and provide them with tantalizing clues to push their investigation along. Even after the reporters published their exposé and Nixon resigned, the journalists kept Deep Throat's identity a mystery. Finally, on April 30, 2005, the *Washington Post* revealed Deep Throat's name: he was Mark Felt, the number two man in the Federal Bureau of Investigation at that time. According to Figure 2.2, Felt acted as a disloyal "leaker." Shortly after the disclosure of Felt's identity, which was done with his consent, he passed away. During his eulogies, he was hailed as a true American patriot.¹³

In the late 1990s, Enron was the envy of Wall Street. It began as a natural gas pipeline company and quickly got involved in the trading of energy-related products in Texas. With its success came access to power. Enron's chair and CEO, Kenneth Lay, was a personal friend of President George H. W. Bush, his chief of staff James Baker, and President George W. Bush. As the company became involved in trading instead of producing, it began to gamble on an unprecedented scale. However, some of its financing was fraudulent, and shady bookkeeping allowed it to hide its true nature from the outside world—that is, until it imploded from within, spurring a congressional investigation into its failure. The congressional hearings revealed that a little-known analyst named Sherron Watkins had written several memos to Kenneth Lay voicing her deep concern over

method of budgeting is unknowable and warned that we should not even try to implement such a system. In sum, he argued that "rational" models deny a fundamental principle of democratic government: the political rationality by which elected representatives balance the numerous and often contradictory concerns of their various constituents. In a democracy, politics is not about efficiency; rather, it is about bargains and compromises among competing interests. Wildavsky dismissed the PPBS because of its single-minded pursuit of economic efficiency—measured in terms of maximum output (benefits) relative to input (costs). Running the government with the efficiency of a private business, he said, was "downright undemocratic and un-American."

While Wildavsky questioned the criteria by which government programs should be evaluated, scholars such as Charles Lindblom and David Cohen pushed to the extreme the argument denouncing quantitative techniques and scoffed at the idea of

the way the company was operating. Her protests were completely ignored. Ms. Watkins quickly came forward and began testifying against her own company and her supervisors, becoming one of the nation's most famous whistleblowers in the process.²⁰

In another dark chapter in American political history, this time in 2004, news began to break that American soldiers were abusing Iraqi detainees in the notorious Abu Ghraib prison. Soon, pictures taken by military police personnel of humiliated, tortured, and abused Iraqi men began to circulate. Freedom from torture is a part of the Universal Declaration of Human Rights and is considered to be the hallmark of any civilized nation. Yet a secret memo, written primarily by then White House lawyer John Yoo, opined that torturing prisoners held in secret U.S. prisons outside its territories was permissible.²¹ He also argued that the Fourth Amendment to the Constitution did not apply to domestic military operations as a part of the larger global war on terror. Such techniques had strong support from most high-level Bush administration officials, including Vice President Dick Cheney and Defense Secretary Donald Rumsfeld.²² Yet none understood the risk of torturing prisoners and abrogating the Geneva Convention better than the Secretary of State Colin Powell. Apart from moral questions surrounding the torture techniques, Powell's military experience allowed him to see the danger that every future American prisoner would face in wars yet to be fought as a result of the practices. He tried to protest forcefully, but lost his own battle within the administration. However, perhaps his military training prevented him from resigning, or he might have reasoned that his presence was an asset as a moderating force. In any case, he remained, continuing to voice his protests behind closed doors.²³

Discussion Points

1. Why do values sometimes conflict within organizations?
2. Compare and contrast how Mark Felt, Sherron Watkins, and Colin Powell dealt with the moral dilemmas they faced.
3. If you were in their place, would you have taken a different action?

using statistics to conduct "professional social inquiry." Instead of quantitative methods, Lindblom and Cohen urged policymakers to use "ordinary knowledge" when making decisions.²⁴ Lindblom declared forcefully, "For all the effort and for all its presumed usefulness, I cannot identify a single social science finding or idea that is undeniably indispensable in any social task or effort. Not even one."²⁵

Lindblom overstates his case and too quickly throws out the proverbial baby with the bathwater by dismissing all quantitative techniques. Peter deLeon notes that policy analysis has made significant contributions to many government programs, such as those created by the 1988 Family Support Act and the 1990 Clean Air Act.²⁶ Besides these highly publicized cases, it is obvious that policymakers—from the chair of the Federal Reserve to the directors of financial management divisions in state and local governments—are influenced by economic analyses conducted in-house or outside their

agencies. They use analytical techniques when projecting revenues and costs, when evaluating the overall desirability of a project, and when submitting an environmental impact statement, which is required for all public projects. They may not always use the most sophisticated techniques of analysis, but they do use systematic, logical thinking, which takes them far beyond the reaches of "ordinary knowledge," far beyond the realm of unstructured, subjective assessment.

However, we should note that quantitative techniques, by themselves, cannot solve all of our problems. The wise people of antiquity told us that although it is important to know what we do know, it is even more important to know what we do not know. Before we discuss the techniques of analysis, we must first describe their limits. These limits arise because

1. we cannot prove or disprove our hypotheses with the precision of laboratory experiments;
2. we cannot eliminate uncertainty;
3. our techniques may offer a limited view of reality;
4. our predilections, prejudices, and biases often cloud our judgment; and
5. we cannot define social welfare.

First, for the most part, hypotheses posed in the natural sciences can be proved or disproved by controlled experiments. If I hypothesize that two hydrogen atoms combined with a single oxygen atom produce water, I can prove my assumption conclusively in a controlled laboratory experiment. This result can then be replicated. However, if a president up for reelection claims that his economic policies caused national prosperity, we have no way of proving the claim (see chapter 5). As a result, eighty years later, liberals credit President Roosevelt's policies for ending the Great Depression, while conservatives remain unconvinced.

Second, in an uncertain world, we make our predictions based on mathematical models. Mathematician John Casti reminds us that regardless of how accurate our measurements and our analytical techniques are, we are never able to completely penetrate the core of uncertainty.²⁷ As you will see in later chapters, there are several techniques designed to aid our decision making in an uncertain situation, but they cannot reach certainty. At the end of one highly mathematical textbook on macroeconomics, a discipline that is useful in analyzing public policy, the author concludes the chapter "Policy Making under Uncertainty" with this startling admission:

Economic analysis has a long way to go until we can specify models with as little residual uncertainty as the ones posited in this chapter, and unfortunately economic analysis has little to say about the appropriate conduct of policy when there is uncertainty as to the correct model of the economy. Therefore, the actual practical usefulness of the analyses of this chapter may be quite limited.²⁸

Third, when we make models of human behavior, we must make assumptions regarding its essence. Since all public policies aim to alter existing patterns of behavior through incentives (tax breaks, monetary rewards, etc.) and disincentives (taxes, fines, quotas, etc.), we must make assumptions regarding human nature. As we have seen in the first chapter, the concept of economic rationality has gathered a lot of scorn from

many quarters. Thus, Alice Rivlin, former president of the American Economic Association and former director of the Office of Management and Budget, took her colleagues to task in the early 1990s for relying too heavily on models: "Economists . . . in their usual fashion, have been short on realism and long on theory and prescription."²⁹ Journalist Anatole Kaletsky even places the blame for the current economic crisis squarely on economists: "The economics profession must bear a lot of the blame for the current crisis. If it is to become useful again it must undergo an intellectual revolution—becoming both broader and more modest."³⁰ Yet the assumption of the "rational" *homo economicus* remains prevalent. Political scientist Elinor Ostrom takes a kinder view of economists and their assumptions about human nature; in a 1997 address to the American Political Science Association, she stated, "While incorrectly confused with a general theory of human behavior, complete rationality models will continue to be used productively by social scientists, including the author."³¹ Echoing Ostrom's sentiment, another political economist asserted that models based on economic rationality are "neither always true nor always good, merely almost always useful."³² Because quantitative techniques are based on a fragmented model of reality, they will always be open to criticism from those who do not agree with the models we have selected. Furthermore, the conclusions we reach using one technique may contradict the results we get using another, or they may vary when we rearrange the data (see chapter 7).

Fourth, as we discussed in the previous chapter, our biases affect our judgments. When we observe, we receive information through filters of culture, religion, and personal life experience. In the social sciences our observations (and the data we gather) admit biases of many kinds. If our data are biased, we do not get an objective result that is accepted by everyone, regardless of how sophisticated our analytical techniques may be (see chapter 6 and the discussion of imperfect data in chapter 9).

Finally, analytical techniques fail us because we cannot define social welfare. Although we are quick to call for the "common good," in reality, it is almost impossible to agree on what is best for society, and the public policies that will achieve it. Every policy creates winners and losers. When we build a new airport, the region as a whole may win through increased trade, commerce, and ease of transportation, but those with buildings near the flight paths lose property value. Spending on education helps the young but takes money from elderly people and those who do not have school-age children. As we will see in chapter 4, developing a satisfactory framework in which to "size up" the winners' wins and the losers' losses may exceed our powers of analysis.

Taking a comfortable middle road, however, will not end the debate over the usefulness of quantitative techniques in public policy analysis. Without a controlled experiment, we cannot prove the effectiveness of our analytical techniques. Even with only "ordinary knowledge," without the aid of "scientific models," we can safely forecast that we will continue to debate this issue—and fill the pages of professional journals—for the foreseeable future. In the meantime, we will use quantitative techniques to analyze public policy. However, we must do so with full knowledge of their strengths and weaknesses. I argue that scientific methods and their claims to objectivity apply only to the relatively narrow process that begins once we have defined the problems, identified the alternatives, and specified the techniques of analysis. *We must make matters of subjective judgment clear at every step of the way, and understand how decisions reflect the goals and principles of public policy analysts and their political bosses.* From observation to analysis,



A CASE IN POINT

Poverty in America: Contrasting Perspectives

After reading the Government Accountability Office report on poverty in Appendix A, you might be persuaded as to the plight of poor people in America, the land of plenty. However, like beauty, nearly every aspect of public policy is based on perception. For instance, the conservative Heritage Foundation published a report on what it means to be poor in America, arguing that most of those whom we call poor are, in fact, not destitute.³³ As evidence, Robert Rector points out the following statistics:

- Forty-three percent of all households identified as poor by the Census Bureau actually own their own homes, which typically have three bedrooms, one and a half bathrooms, a garage, and a porch or patio.
- Eighty percent of these homes have air conditioning.
- About 6 percent of the homes are "overcrowded," and more than two-thirds have more than two rooms per person.
- Compared with the average (middle-class) European living in Paris, London, Vienna, Athens, and other cities, poor people in America enjoy larger accommodations.
- Almost 93 percent of poor Americans own cars, with 31 percent owning two cars.
- Over 97 percent of poor households in American have color televisions, with half having two.
- Similarly, most of these "poor" households have microwave ovens, stereos, automatic dishwashers, DVD players, and cable or satellite television.

Rector concludes, "The living conditions of persons defined as poor by the government bear little resemblance to notions of 'poverty' promoted by politicians and political activists." For him,

the most wonderfully complex computers in the world guide us: our brains. The quality of our analysis will be judged in the end not by the sophistication of our technique but by the wisdom of our judgment.

STRUCTURE ABOVE A SWAMP: PUBLIC POLICY ANALYSIS AND PROFESSIONAL ETHICS

Now you may be confused. If everything about public policy is so vague and subjective, what is the use of learning methods of objective research? On one hand, I emphasize the relative nature of truth, and on the other, I advocate objective analysis. A quotation from Karl Popper, the eminent philosopher of science, may resolve this contradiction:

The empirical basis of objective science has thus nothing "absolute" about it. Science does not rest upon rock-bottom. The bold structure of its theories rises, as it were, above a swamp. It is like a building erected on piles. The piles are driven down from above into the swamp, but not down to any natural or "given" base; and when we cease our attempts to

poverty, when it does exist, is caused by the fact that "the typical poor family with children is supported by only 800 hours of work during a year: That amounts to 16 hours of work per week." He identifies children born out of wedlock as another major contributing factor. Therefore, Rector recommends that any antipoverty measure must include incentives for mothers to work longer hours as a condition for receiving aid and that welfare and tax laws should be amended to promote marriage.

Ron Haskins and Isabel Sawhill of the liberal Brookings Institution, using the same set of information, come to a radically different conclusion, and consequently, their policy prescriptions are different from those of Rector. They see inequality in the United States as growing rapidly and recommend the following³⁴:

- Strengthen work requirements in government assistance programs.
- Increase the minimum wage.
- Expand the earned-income tax credit.
- Subsidize child care for low-wage workers.
- Promote marriage as the best environment for rearing children.
- Fund effective teen pregnancy prevention efforts, premarital education, and family planning.
- Invest in high-quality early childhood education.

Discussion Points

1. How do differences in values affect an analyst's judgment?
2. Discuss how the two perspectives on poverty stem from the respective ideological stances of the two reports' authors.
3. If you were working as a policy analyst, what would you do to avoid an ideological bias?

drive our piles into a deeper layer, it is not because we have reached firm ground. We simply stop when we are satisfied that they are firm enough to carry the structure, at least for the time being.³⁵

Therefore, although we can empirically test "scientific laws" regarding society, we cannot know their truth. The ability to test hypotheses has lent social sciences and policy science a considerable degree of credibility. Hence, on this shifting ground of "truth," we want to achieve objective analyses by being systematic in our definitions of goals, consistent about our methods of analysis, and forthright about our implicit assumptions.

In building castles over shifting sand, we must rely not only on the appropriateness of the tools but also on the integrity of those who must use them. The ethics of public policy is an important area of consideration. In this discussion, I can point to two broad areas of bias: personal honesty and the recognition of ideological biases.

At the personal level, analysts must overcome a number of ways in which their positions of power and knowledge can make them vulnerable to corruption. Analysts may

often face a **conflict of interest**. One of the most difficult conflict-of-interest situations to grapple with—since it may happen under the radar and go unnoticed—is **honest graft**. Suppose I am employed by the city planning department and am working on a proposal for the expansion of mass transit to an area that is currently known for urban blight. I know that when the project is implemented and the new rapid transit system is introduced, the land value will increase substantially. Armed with this privileged knowledge, I choose to invest in the area that will benefit the most from the ensuing economic boom. If I do that while I am preparing the report, my analyses and my recommendations may be influenced by my own selfish interest. I may want to emphasize the benefits and downplay the costs.

To take another example, say that during your work as a policy analyst, you have developed specialized knowledge that may be of interest to a private corporation. You believe that the corporation may be willing to offer you a high-paying job. If you are writing a report that has a significant impact on the company, you may be tempted to alter your report to make your policy recommendations favor your future employer. If you want to be totally impartial, in such a situation you must make full financial disclosure and, under certain circumstances, may recuse yourself from the project.

Your professional integrity may also be compromised if you orient your report to fit with the ideological or political agenda of your superiors. In most federal agencies, the party in power appoints the top management. Therefore, for career advancement, you may be tempted to slant your analysis to support your ideologically driven boss.

Finally, you may be driven by your own ideological position. In that case, you may deliberately ignore evidence that is contrary to your value position and cherry-pick information in a way that supports your view. Once again, it is exceedingly difficult to completely eliminate the biases arising out of your own ethical norms, since you may not even recognize it as an **ideological bias**.

In an otherwise subjective world, riven by prejudiced assessment, skewed perception, and the influence of personal interest, how do we maintain objectivity? I argue that these corrupting influences can be addressed only when analysts take pride in their profession. Their professional integrity is the guiding light that, at the end of the day, must ensure that their analyses are protected from systematic biases. In this context, perhaps, the work of prominent philosophers can provide an answer. We may recall John Stuart Mill's admonition of doing no harm as a moral compass. To this we may add John Rawls's (see chapter 5) assertion that a society should be judged by the welfare of those who are at the bottom. Thus, in evaluating public policies we must make sure that the burden of cost does not fall primarily on those who can least afford it.

In this book we will learn a number of different statistical techniques that are used for policy analyses. These tools carry with them an aura of objectivity, which can be extremely misleading. In their scientific garb, these "objective" analyses can hide strong ideological biases. In psychology it is often held that the strength of one's character can also be the source of one's weakness. Similarly, the appeal of objective methods of policy analysis is the ability to present complex phenomena with simple, easy-to-understand numbers and figures. At the same time, the unquestioned acceptance of these statistical approaches can lead to serious flaws. Therefore, we should know how to use with skill the extremely

useful and powerful tool called statistics. This skill is honed with practice and by knowledge of the methods of manipulation. As we embark upon learning the tools of the trade for policy analysis, we may begin where John Maynard Keynes ended his seminal 1936 book, *The General Theory of Employment, Interest and Money*. He correctly warned,

The ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly believed. Indeed, the world is ruled by little else. Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back. I am sure that the power of vested interests is vastly exaggerated compared with the gradual encroachment of ideas. Soon or late, it is ideas, not vested interests, which are dangerous for good or evil.³⁶

As the profile of policy analysis has risen in recent years, so has the volume of literature covering the hard issues of professional ethics. Echoing their colleague Mark Lilla,³⁷ social scientists David Weimer and Aidan Vining suggest that the profession of policy analysis should work toward a new standard of conduct that “explicitly recognize[s] our obligations to protect the basic rights of others, to support our democratic processes as expressed in our constitutions, and to promote analytical and personal integrity.”³⁸ As an analyst attempting a high-wire act, balancing conflicting values, interests, and concerns, you can do no better than observe this standard.

KEY WORDS

Coequal model (p. 31)

Conflict of interest (p. 42)

Honest graft (p. 42)

Ideological bias (p. 42)

Mixture-in-administration model (p. 31)

Mixture-in-policy model (p. 32)

Policy-administration dichotomy model (p. 30)

Rational model (p. 35)

Whistleblowers (p. 34)

EXERCISES

1. Write an essay on the role of a policy analyst in a public agency. Discuss the various balancing acts this analyst must perform.
2. There is a continuing debate on the use of statistical and operations research techniques in the public policy making process. What are the strengths and weaknesses of these techniques?
3. Should policy analysts be dispassionate scientists, or should they express their deeply held values? Consider an important public policy issue that your community is currently debating. How would you analyze the issue? Do you think your analysis would be influenced by your values, your ideology, or the organization you work for? If so, should your analysis be considered “objective”?

Notes

1. For a delightful discussion, see Peter deLeon, "Reinventing the Policy Sciences: Three Steps Back to the Future," *Policy Sciences* 27 (March 1994): 77–95. Also see Herbert Goldhamer, *The Adviser* (New York: Elsevier, 1978), and Arnold J. Meltsner, *Rules for Rulers: The Politics of Advice* (Philadelphia: Temple University Press, 1990).
2. Although it is a bit dated, for an excellent discussion, see Arnold J. Meltsner, *Policy Analysts in the Bureaucracy* (Berkeley: University of California Press, 1976). Also see Robert Formaini, *The Myth of Scientific Public Policy* (New Brunswick, N.J.: Transaction, 1990), and Robert A. Heineman, *The World of the Policy Analyst: Rationality, Values, and Politics*, 2nd ed. (Chatham, N.J.: Chatham House, 1997).
3. See Allan M. Brandt, *The Cigarette Century: The Rise, Fall, and Deadly Persistence of the Product That Defined America* (New York: Basic Books, 2007).
4. Woodrow Wilson, "The Study of Administration," *Political Science Quarterly* 2 (1887): 197–222.
5. The articles by the "rebel academics" were published in *Toward a New Public Administration: The Minnowbrook Perspective*, ed. Frank Marini (San Francisco: Chandler, 1971).
6. Frederick C. Thayer, *An End to Hierarchy! An End to Competition! Organizing the Politics and Economics of Survival* (New York: New Viewpoints, 1973); later republished as *An End to Hierarchy and Competition: Administration in the Post-affluent world* (New York: New Viewpoints, 1981).
7. Lewis C. Mainzer, *Political Bureaucracy* (Glenview, Ill.: Scott, Foresman, 1973).
8. Svara, "Dichotomy and Duality."
9. In 1514, Henry requested that Pope Leo X perform the highly unusual but not entirely unprecedented action of making Wolsey (then a bishop) a "Cardinal Sole." In his letter to the Pope on August 12, Henry claimed that he "esteemed Wolsey above his dearest friends and could do nothing of importance without him." Jasper Ridley, *Statesman and Saint: Cardinal Wolsey, Sir Thomas More, and the Politics of Henry VIII* (New York: Viking, 1982), 48.
10. Henry showed deep personal respect and friendship for More and often stated that he "valued no one's approval more than his." See Carolly Erickson, *Great Harry: The Extravagant Life of Henry VIII* (New York: Summit, 1980), 120.
11. Albert O. Hirschman, *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States* (Cambridge, Mass.: Harvard University Press, 1970).
12. David L. Weimer and Aidan R. Vining, *Policy Analysis: Concepts and Practice*, 3rd ed. (Upper Saddle River, N.J.: Prentice-Hall, 1999).
13. See, for example, Robert S. Lynd, *Knowledge for What? The Place of Social Science in the American Culture* (Princeton, N.J.: Princeton University Press, 1939); Harold Lasswell, *Power and Personality* (New York: Norton, 1949); and Daniel Lerner and Harold Lasswell, eds., *The Policy Sciences* (Palo Alto, Calif.: Stanford University Press, 1951).
14. Frank Fischer, *Evaluating Public Policy* (Chicago: Nelson-Hall, 1995), chap. 1.
15. For a comprehensive look at the field of operations research and quantitative methods of decision making in the area of public policy analysis, see Göktuğ Morçöl, ed., *Handbook of Decision Making* (Boca Raton, Fla.: CRC, 2007).
16. See, for example, Virginia Held, "PPBS Comes to Washington," in *Planning-Programming-Budgeting: A Systems Approach to Management*, ed. Freemont J. Lyden and Ernest G. Miller (Chicago: Markham, 1968), 11–26. For one of the earliest explanations of "systems theory" in politics, see David Easton, "An Approach to the Analysis of Political Systems," *World Politics* 9 (1957): 383–400.
17. For a discussion of President Reagan's executive order and the use of quantitative techniques, see William Ascher, "The Evolution of Policy Sciences," *Journal of Policy Analysis and*

- Management* 5 (1986): 365–389; and Peter deLeon, *Advise and Consent: The Development of the Policy Sciences* (New York: Russell Sage Foundation, 1988). Also see Douglas Torgerson, "Between Knowledge and Politics: Three Faces of Policy Analysis," *Policy Sciences* 19 (1986): 33–60.
18. Aaron Wildavsky, *The Politics of the Budgetary Process*, 3rd ed. (Boston: Little, Brown, 1979), 197.
 19. David von Drehle, "FBI's No. 2 Was 'Deep Throat': Mark Felt Ends 30-year Mystery of *The Post's* Watergate Source," *Washington Post*, June 1, 2005.
 20. For an excellent discussion, see Bethany McLean and Peter Elkind, *The Smartest Guys in the Room: The Amazing Rise and Scandalous Fall of Enron* (New York: Portfolio Trade, 2004).
 21. For an excellent discussion of moral debate on the issue, see Karen J. Greenberg, ed., *The Torture Debate in America* (Cambridge, Mass.: Harvard University Press, 2006).
 22. For a real-life experience of an interrogator on the effectiveness of torture as a tool for extracting information from accused terrorists, see Matthew Alexander (with John R. Bruning), *How to Break a Terrorist: The U.S. Interrogators Who Used Brains, Not Brutality, to Take Down the Deadliest Man in Iraq* (New York: Free Press, 2008).
 23. Jane Mayer, *The Dark Side: The Inside Story of How the War on Terror Turned into a War on America's Ideals* (New York: Doubleday, 2008), 292.
 24. Charles E. Lindblom and David K. Cohen, *Usable Knowledge: Social Science and Social Problem Solving* (New Haven, Conn.: Yale University Press, 1979).
 25. Charles Lindblom, *Inquiry and Change: The Troubled Attempt to Understand and Shape Society* (New Haven, Conn.: Yale University Press, 1990), 131.
 26. deLeon, "Reinventing the Policy Sciences," 78.
 27. John L. Casti, *Searching for Certainty: What Scientists Can Know about the Future* (New York: Morrow, 1990).
 28. Stephen McCafferty, *Macroeconomic Theory* (New York: Harper & Row, 1990), 360–361.
 29. Alice M. Rivlin, "A New Vision of American Federalism," *Public Administration Review* 52 (1992): 315–320.
 30. Anatole Kaletsky, "Goodbye, Homo Economicus," *Prospect*, April 26, 2009.
 31. Elinor Ostrom, "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997," *American Political Science Review* 92 (March 1998): 1–22.
 32. Mark I. Lichbach, *The Rebel's Dilemma* (Ann Arbor: University of Michigan Press, 1995), 344.
 33. See Robert E. Rector, "How Poor Are America's Poor? Examining the 'Plague' of Poverty in America," The Heritage Foundation, August 27, 2007, www.heritage.org/research/welfare/bg2064.cfm.
 34. Ron Haskins and Isabel V. Sawhill, "Attacking Poverty and Inequality: Reinvigorate the Fight for Greater Opportunity," Brookings Institution, 2008, www.brookings.edu/papers/2007/~/media/Files/Projects/Opportunity08/PB_Poverty_Haskins_Sawhill2.pdf.
 35. Karl Popper, *The Logic of Scientific Discovery* (New York: Harper & Row, 1959), 65.
 36. John Maynard Keynes, *The General Theory of Employment, Interest and Money* (New York: Harcourt, Brace, 1936).
 37. Mark T. Lilla, "Ethos, 'Ethics,' and Public Policy," *Public Interest* 63 (1981): 3–17.
 38. Weimer and Vining, *Policy Analysis*, 57.

3

Government and the Market

The city of Boston and its surrounding region faced a housing crisis in the 1990s.¹ As the region experienced a nearly decade-long economic boom, jobs were created and new housing developments sprang up everywhere. However, housing was priced far beyond the reach of low-income residents. According to a government study, only eleven communities in the greater Boston area set aside more than 8 percent of their housing for low-income families. Ninety-three communities failed to do even that much. After months of pressuring reluctant suburban politicians, Mayor Thomas Menino submitted legislation to the city council to penalize neighboring cities if they did not set aside at least a tenth of their housing in the affordable price range. If they did not meet this goal, and made no effort to meet it, the legislation recommended that the state would withhold housing funds. Today, many cities across the land, particularly those with very high housing costs, have passed similar ordinances. This example may cause you to wonder why government interferes in a free market system. If some of its citizens cannot afford to live in a particular city, is it the proper role of the government to force the city to create housing for them? If so, on what basis should it intervene?

History tells us that governments have always strongly influenced the lives of citizens. In fact, until recently, the grip of the king and the church was quite strong. Protests against such domination started in Europe. In 1776, as the first shots of rebellion were being heard in the colonial states, a revolution of a different sort was taking place on the Atlantic's other shore. Adam Smith, an ex-customs official and a tutor to an aristocratic child, published *An Inquiry into the Nature and the Causes of the Wealth of Nations*.² As noted in chapter 1, Smith wrote his book to repudiate the then popular theory of **mercantilism**, which argued that a nation becomes prosperous when it maintains a trade surplus by maximizing exports and minimizing imports. To limit imports, mercantilism erected trade barriers. In contrast, Smith

argued that the wealth of a nation increases when all restrictions to the free functioning of the market are removed. He reasoned that individuals, when free of outside restrictions, pursue their self-interest, allowing the nation to prosper.

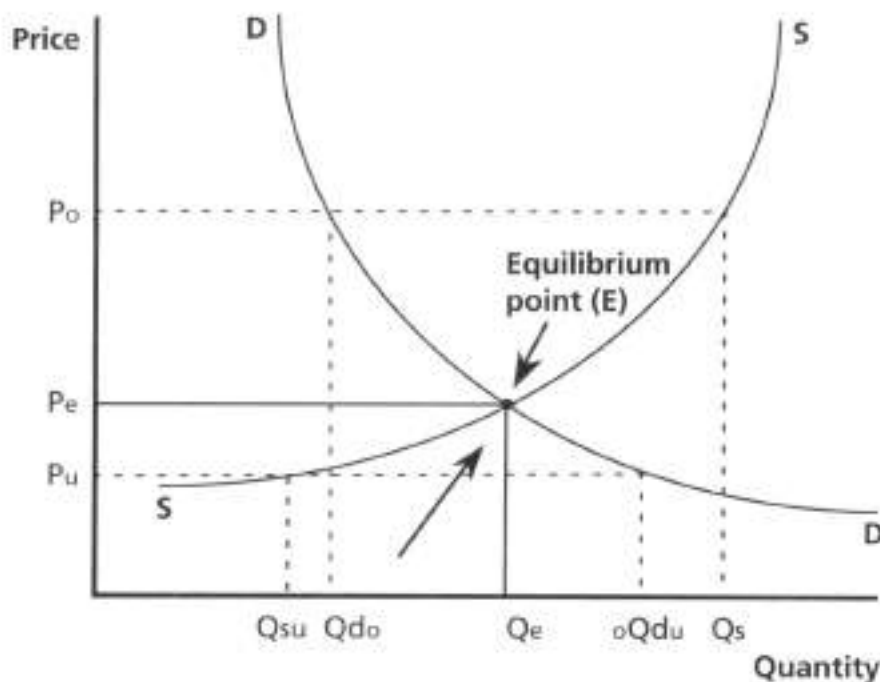
Adam Smith's arguments regarding the efficient working of a market are shown in Figure 3.1.³ The curve DD depicts the demand for a certain commodity, while the SS curve traces its supply. The point of **equilibrium** is reached where the curves intersect. This point is called *equilibrium* (a term borrowed from mechanics) for two reasons:

1. When the market rests at that point, no internal force will cause it to move away from equilibrium.
2. If the market departs from this point, the push and pull of demand and supply will bring it back to equilibrium (point E).

If a producer asks for a price higher than is warranted by the market (P_o), the demand for the product will dwindle, and ultimately the seller will have to lower the price. Similarly, if one producer makes an abnormally high profit by selling a new product, soon other suppliers will flood the market with copycat products. A situation of oversupply will occur, pushing the price down from P_o to P_e . The opposite scenario will take place if the market price is below the equilibrium point. Consumers will attempt to purchase more goods while suppliers cut down on production. This situation of excess demand drives the market price from P_u to P_e .

FIGURE 3.1

The Foundation of Economic Reasoning: Market Equilibrium



Note: The demand curve, DD, is downward sloping because as the price of a commodity goes down, its demand increases. In contrast, the supply curve, SS, slopes upward. When prices go up, supply increases in response to a greater expectation of profit. The point of intersection of the two curves is called the equilibrium point.

This diagram—perhaps the most well-known illustration in the social sciences—forms the foundation of the entire field of economics. As you can see, if there is any outside intervention, the market is corrupted; it becomes inefficient because it cannot reach the point of equilibrium. When supply is arbitrarily restricted, buyers pay a higher price than they should be paying. This situation leads to unemployment and other market inefficiencies. On the other hand, when supply is increased beyond the point of equilibrium, buyers consume more than they should, leading to overuse and waste.

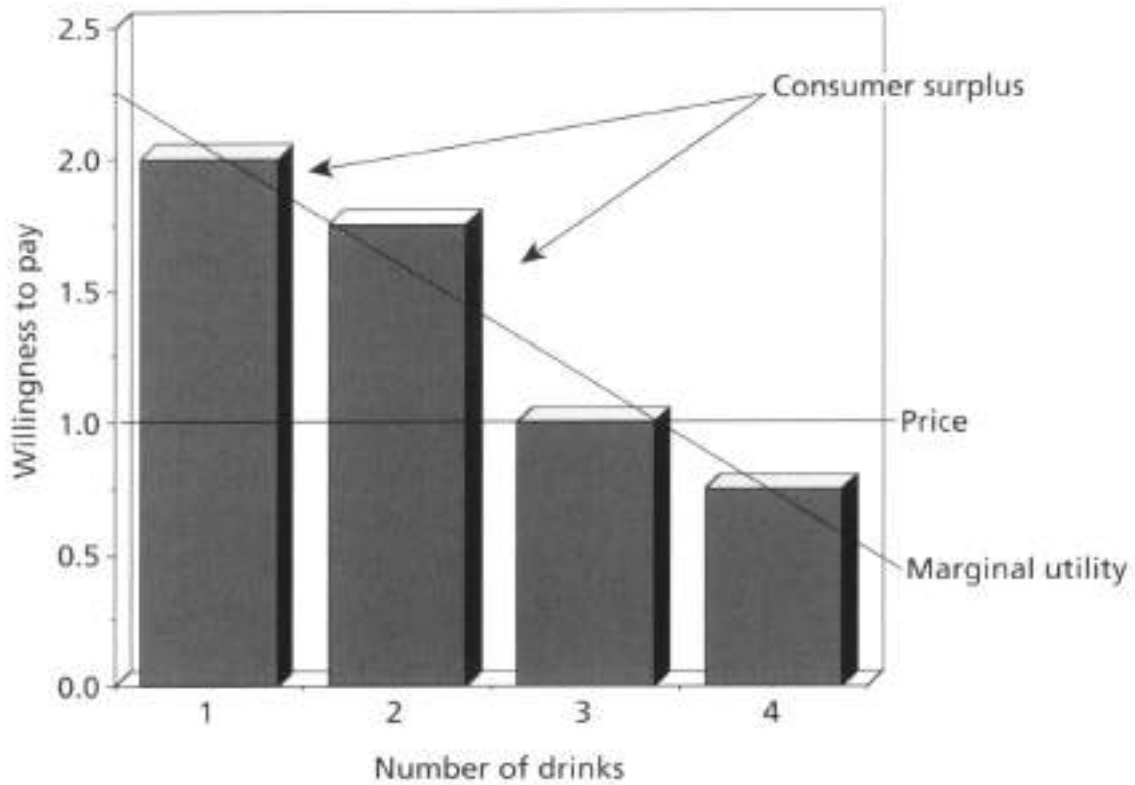
You may note the normative implications of this discussion, which are apparent from our use of the imperative *should*. The government *should* stay away from the marketplace so that prices reflect the “true” forces of demand and supply. When the government does not interfere, the invisible hand of market forces clears the market and allocates resources among the four factors of production—land, labor, capital, and organization—in the fairest way. Landowners get the rent they deserve, and laborers get the wages they merit. Moreover, market forces allocate the most appropriate amount of interest to the owners of capital and the right level of profits to the entrepreneurs. When anyone is paid more than he or she deserves, the market wipes out the excess profit through free competition.

THE FUNDAMENTAL CONTRIBUTIONS OF ECONOMIC ANALYSES

The success of an academic discipline can ultimately be gauged by the extent to which its core doctrines permeate our daily lives, social values, and accepted norms. By that measure, the intellectual revolution sparked by the bored Scottish tutor of a nobleman's son is unparalleled in social science. In his analysis of a nation's prosperity, Smith placed the individual at the forefront of the economic system. Over the 200 years following the publication of *Wealth of Nations*, economics has grown in many diverse directions. If we were to point out the most fundamental contributions of economics as an analytical tool, we might mention the related concepts of **marginal analysis** and **opportunity costs**. These concepts are so integral to systematic analysis that it is virtually impossible to engage in a meaningful academic discourse about public policy without taking them into account.

Marginal Analysis

In the economic literature the term *marginal* means “additional.” Although the word is used in classrooms as well as boardrooms, policymakers often make serious mistakes when they confuse *marginal* with *average*. Let me explain marginal utility by giving you an example. Suppose you are thirsty and you come upon a stand selling cold drinks. You are so thirsty that you are willing to pay \$2 for the first drink. However, the drink is too small to quench your thirst, so you want a second one. Because you are already somewhat satisfied, you are willing to spend up to \$1.75 for it. After quaffing the second drink, you are still thirsty. You are willing to pay \$1 for a third and \$0.50 for a fourth. We plot your willingness to pay in Figure 3.2. Remember, these numbers reflect only

FIGURE 3.2 Marginal Utility and Consumer Surplus

your psychological willingness to pay, not the actual prices for the marginal, or additional, drinks. They represent the marginal utility you receive from the drinks.

The important question becomes, at what point would you stop buying drinks? Let us say the price of a drink is \$1. You will stop after you have consumed three drinks. You will not stop before then, because you are getting a bargain with each drink, and each provides you with more utility than it costs you. However, the price of the fourth drink is below your marginal utility. The first rule of **consumption efficiency** is to *stop consuming at the point where the marginal utility of the good is exactly equal to its marginal cost*. Your marginal cost (the cost of an additional drink) is fixed at \$1. In a perfectly competitive market, you are a **price taker**. That is, because the market is too large for you to manipulate, you cannot influence the market price with your purchase. The additional utility you derive from three drinks, the amount above your market price (\$1), is your **consumer surplus**. It is the difference between the market price and the total utility. For the first drink, the consumer surplus is \$1 (\$2 – \$1); for the second drink, it is \$0.75 (\$1.75 – \$1); and for the third drink, there is no consumer surplus (\$1 – \$1). The full implications of this concept are discussed in chapter 14.

So as a consumer, you know when to stop buying. What about the producer? A producer produces to the point where his or her **marginal revenue** equals the price. Suppose I am the owner of that drink stand and I am deciding how many gallons of drinks to produce. As I sell each drink, I make a certain amount of revenue. I keep buying the ingredients until I reach the point where the marginal cost (of buying, transporting, storing, and serving) is equal to the marginal revenue from the product. Because I am a

small producer and cannot influence prices in the market, I am—like the customer—a price taker. Therefore, *I will stop producing when my marginal revenue is equal to the marginal cost.* Together, the two dictums about consumption and production constitute the iron laws of microeconomics. They give us both the point at which a consumer would stop buying and the one at which a producer would set limits for production.

Marginal Versus Average: Cause for Policy Confusion. Although marginal costs and benefits are basic concepts of microeconomics, they are often misunderstood. Two examples, one drawn from the private sector and the other from a public policy debate, can help us understand the confusion.

The next time you board an airplane, look around the aircraft. Although each passenger will travel the same distance, the fares they pay will vary widely. Some passengers have paid full fare for, say, buying the ticket after the minimum number of stipulated days or for not spending a Saturday night at their destinations. There are those who bought tickets at a discount due to their age (senior citizens or children), special promotional deals, or special arrangements with travel agencies. Besides these passengers, there may be students on board with cut-rate standby fares or those who are flying free of charge on a “frequent flyer” arrangement. Does this wide variation surprise you? Does the airline lose money for selling certain tickets at a discount, since each discounted ticket brings down the average price per ticket?

The surprising answer is that the airline makes money even when it sells tickets at a significant discount (awarding free tickets to frequent flyers may fall under the separate category of giving incentives or promoting passenger loyalty). This is because most of the costs of flying an airplane are fixed. The salaries of the crews and the ground staff must be paid regardless of the number of passengers on board. The port authorities must be paid for landing rights at the airport. The costs of maintenance will not vary significantly with the *additional* number of passengers. Therefore, the marginal cost of carrying an additional passenger is negligible. After those paying full price have covered maintenance costs, any revenue received from a senior citizen or a standby passenger will result in positive marginal revenue for the airline. If decision makers look only at the average fare—which goes down with every discounted ticket—they will surely miss the opportunity to increase the company’s profit by carefully instituting discounted prices.

When it comes to the public sector, the confusion about average and marginal costs can lead to serious policy mistakes. Take, for example, the controversy over the cost of illegal immigration in California. During the mid-1990s, the state was caught up in a divisive debate over the cost of educating the children of illegal immigrants. A referendum was passed barring these children from public education. During the debate the proponents of the ban estimated the costs by taking the *average cost* per child (the entire state education budget divided by the number of students) and multiplying it by the number of children of illegal immigrants enrolled in California’s public schools. In contrast, those who opposed the measure pointed out that much of the education budget is fixed. They argued that if we want to calculate the true cost of educating these children, we must look at the *marginal cost*. It is interesting to note that the two sides were arguing past one another without realizing the sources of the two widely varying figures.

Opportunity Costs

To understand opportunity costs, let us go back to the case of affordable housing in the Boston area. The state of Massachusetts set aside \$400 million in housing aid as an incentive for cities to supply affordable housing to low-income families.⁴ The mayor's plan also called for using some of the state's surplus funds to supplement this housing trust fund. Of course, it is a noble idea to encourage cities to build houses for low-income people. But the money spent on affordable housing projects could have been put to other uses, such as making the streets safer, providing children with a better education, or assisting elderly people and those who are homeless. This argument is normal in public debate; when we spend money for a project, other deserving projects remain unfunded or underfunded. When considering a project, it is common to take into account the costs of alternative projects. In other words, *what must you give up to get the project you are buying?* The answer to this question allows you to determine the opportunity cost of an activity.

Let us take a look at the cost of higher education. How much are you paying for your college education? If you are a full-time student, you would calculate your expenses for tuition, room and board, supplies, transportation, and so on. In addition, you must factor in another cost. You gave up a job you otherwise would have had, had you not attended college. We must add the monetary costs and the estimated forgone income. It is impossible to overlook opportunity cost in public policy decisions, because no resource is infinitely abundant and without alternative uses. Space, clean air, and ocean water all have alternative uses and opportunity costs associated with their exploitation. Harvesting an old-growth forest in the Pacific Northwest clashes with saving the endangered spotted owls. The sudden popularity of fishing for orange roughy off the Australian coast has deprived giant squid of their primary food source. A proposed ballpark vies for public money with other city needs, such as library facilities, garbage collection services, and programs to care for the downtown homeless population. When it comes to public policy analysis, economics—as a powerful analytical tool—teaches us to look beyond the market value of a resource to its opportunity cost.

MARKET FAILURE: WHY GOVERNMENT INTERFERES IN A FREE MARKET

Adam Smith saw the market as inherently self-correcting. If only we let the market run its course, he said, most of our miseries would be solved—if not immediately, certainly in the long run. Yet his critics pointed out that not all markets are self-correcting. The problems of market failure became painfully clear in 1929, when, following the stock market crash, the Great Depression set in. John Maynard Keynes argued that the free market, if left on its own, would not be able to extricate itself from the quagmire of high unemployment (recession or depression) or high prices (inflation). For the economy to reach full employment with stable prices, it would require active government intervention. Furthermore, Keynes famously commented that the market may indeed correct itself in the long run—a favorite term of conservative economists—but “in the long run we are all dead.”

Under the guidance of Keynes, the role of the government in the domestic and international economies became firmly rooted. Building on his work, Keynes's intellectual

descendants determined that the government has four major roles: *allocation, distribution, stabilization, and growth*. First, because the market cannot allocate public goods (see chapter 1), the government must step in. For example, the government builds infrastructure, mandates clean air, and provides public education and military protection because the market fails to supply them. Second, because the market inherently favors rich and powerful people, the government must help those who are caught in an endless cycle of poverty. In this way it plays a redistributive role, whereby through taxation, similar to the mythical Robin Hood, it takes money away from the wealthy and spends it on the poor. Third, with active government intervention through stabilization policies, the market is protected from wild swings of inflation and recession and safeguarded against the extremes of business cycles. Finally, because the market is a poor allocator of resources between current consumption and future investment, an activist government—through its policies that support research and development in science, technology, and education—can ensure future prosperity for the nation.

A society based on free market principles and requiring no government intervention—except to maintain external and internal security and a fair judicial system—sounds good in theory. Upon close scrutiny, however, the society that self-corrects and attains equilibrium loses some of its appeal. When we realize that a perfectly competitive market depends on some of the most restrictive assumptions and truncated views of reality, it is apparent that the market will not always self-correct. A free market is based on four drastic assumptions:

1. There are numerous buyers and sellers.
2. There is perfect exit from and entry into the market.
3. There is a perfect flow of information.
4. There is no **externality**, or no one is affected by the activities of any other person.

A perfectly competitive market is shattered when any of these four conditions is violated. Let us examine these conditions in more detail.

Lack of Competition

The competitive market that emerged from Adam Smith's analyses was rooted in the early days of the industrial revolution, when there were no significant monopoly forces in the market. When there are numerous buyers and sellers, no individual or firm can single-handedly influence the outcome of the market. Each participant in the market is a price taker, as opposed to being a monopolistic **price maker**. Every buyer and seller must adjust his or her respective demand and supply in view of prevailing market prices. Landowners, laborers, moneylenders, and business entrepreneurs earn according to their respective merit. If anyone makes more than that, the excess profit will bring in other competitors, thereby eliminating any profit over what is warranted by merit.

Market conditions have changed significantly since the eighteenth century. Today the market is often plagued by monopoly power. The federal government's antitrust case against Microsoft is an example. The government argued that by bundling its Internet Explorer Web browser with its Windows operating system, Microsoft made it hard for consumers to use another browser, such as Netscape. In its price-making capacity, a monopolist sets a price that guarantees a profit over what it would have earned if there



A CASE IN POINT

Too Big to Fail: How Monopolies Can Hold the Government Hostage

I have explained the theory behind the hypothesis that a monopoly is inherently less efficient than a perfectly competitive market. However, with its economic power a monopoly also acquires political power and positions of special privilege, which can be squarely in conflict with the national well-being. The abuse of power from corporate monopoly became apparent as the nation and the entire world were industrializing. And the most essential ingredient in the process was oil. Sensing an opportunity, the Standard Oil Company brought it under the control of a single entity. Through a series of innovative and stealthy business practices in the 1870s, the legendary John D. Rockefeller, his brother William, and a handful of other men created a huge monopoly by forcing out the competition. They created a "trust," which was a conglomeration of a number of powerful companies. This type of economic giant quickly became known for its predatory practices, leading Congress to pass sweeping antitrust legislation, including the Sherman Act of 1890 and the Clayton Act of 1914.⁵ While they were minimally able to curb the power of the monopolies in the area of oil production and distribution, monopoly thrived in other areas.

Of course, monopoly is not restricted to industry; it involves labor unions as well. As the monopoly power of the United Auto Workers increased, it forced car manufacturers (particularly General Motors and Chrysler) to grant its workers salaries and benefits that were unsustainable in the face of international competition. The rapidly increasing gas prices and environmental concerns associated with large trucks and SUVs, however, soon caught up with GM; it saw its sales plummet and faced the ominous prospect of bankruptcy when the recession of 2008 took hold. Yet GM's going out of business was unthinkable for the policymakers in Washington, leading them to pass a \$14 billion bailout for the automaker. Not only did the automakers employ a

were competition. In 1999, the court decided that in a competitive market, Microsoft would have charged \$29 for Windows 98, but because it was a monopolist, it could charge \$49, nearly double the competitive amount.⁹

A market can sustain a higher price only when the supply is restricted; to hold a high price, a monopolist must reduce the total supply in the market. In a monopolistic market, society loses because of the restricted supply. In addition, a monopolist can choke off innovations because it has little financial incentive to make improvements for a captive group of buyers.

A monopolist does more than simply restrict its output to raise the prices of its products. Some economists have argued that because a monopolist has little incentive to lower costs to achieve maximum efficiency, it often operates at a less than optimal level of efficiency. Harvey Leibenstein coined the term *x-inefficiency* to describe such a situation.¹⁰ In the late 1980s the Defense Department found itself in a public relations controversy when someone pointed out that it was purchasing items such as toilet seats and screwdrivers at exorbitant prices from its monopolistic suppliers. In summer 2000

huge number of people in the nation, they were highly concentrated in certain areas, such as Michigan, where their collapse would have led to an unmitigated disaster.⁶

The American International Group is the largest insurance company on the planet, the eighteenth largest company of any kind in the world. It commanded \$1 trillion in reported assets, roughly equal to the gross domestic income of India,⁷ and it is a component of the Dow Jones Industrial Average. AIG began to experience problems, however, when one of its subsidiaries wrote insurance protection against losses on debts and loans of borrowers to the tune of \$447 billion, spurring a chain of disastrous events. When the housing market was hit by a flood of mortgage defaults and debt piled up in 2008, AIG was hit hard and was left with a very low level of capital reserves. The government stepped in to prevent its total collapse, bailing it out for an eye-popping \$173.3 billion. In return, taxpayers received an 80 percent stake in AIG.

How did these corporations get to be too big to fail? Republicans and Democrats alike had paid scant attention to the problems that monopoly creates for the market. President George W. Bush's Justice Department did not file a single case against a dominant firm for violating the antimonopoly laws.⁸ At the time of this writing, the Obama administration's new policy more closely aligns the American antitrust policy on monopolies and predatory practices with the views of antitrust regulators at the European Commission. If history is any predictor, however, as soon as the economy turns around and public attention shifts, the economic power of the corporate giants will ensure that the concentration of economic power will continue to grow unabated.

Discussion Points

1. Should we be concerned about monopoly power in the current economic market? Why or why not?
2. How can we ensure that the government enforces antitrust laws when public focus on the ills of concentration of economic power shifts?

the oligopolistic oil industry was accused of price gouging. In response, President Bill Clinton, as a stopgap measure, decided to release some of the strategic oil reserves held by the government for such an emergencies.¹¹ (See "A Case in Point" for a more in-depth discussion of monopolies and government actions.)

Barriers to Entry and Exit

The concept of ease of entry into the market has three aspects. The first is **factor mobility**. Take, for example, unemployment. In a perfectly competitive market, it does not exist. This is because an unemployed person is one who is able and willing to accept a job at the going wage.¹² If I want to become a guard in the National Basketball Association but lack the necessary talent and physical skills, I cannot be classified as unemployed, because I am willing to play the position. Similarly, if I decline to take a job at the going rate because it is far below what I expect to earn, I am not considered unemployed, because I can always take this job. If wages are too high, the demand for labor

goes down and exerts a downward pressure on the wage rate. In contrast, if wages are too low, firms hire more people until the market wage rate becomes equal to the market-clearing equilibrium rate. This typically has been the conservative argument in the United States against minimum wage restrictions; if there is unemployment it is because the government has set a minimum wage rate, which is a barrier to hiring.

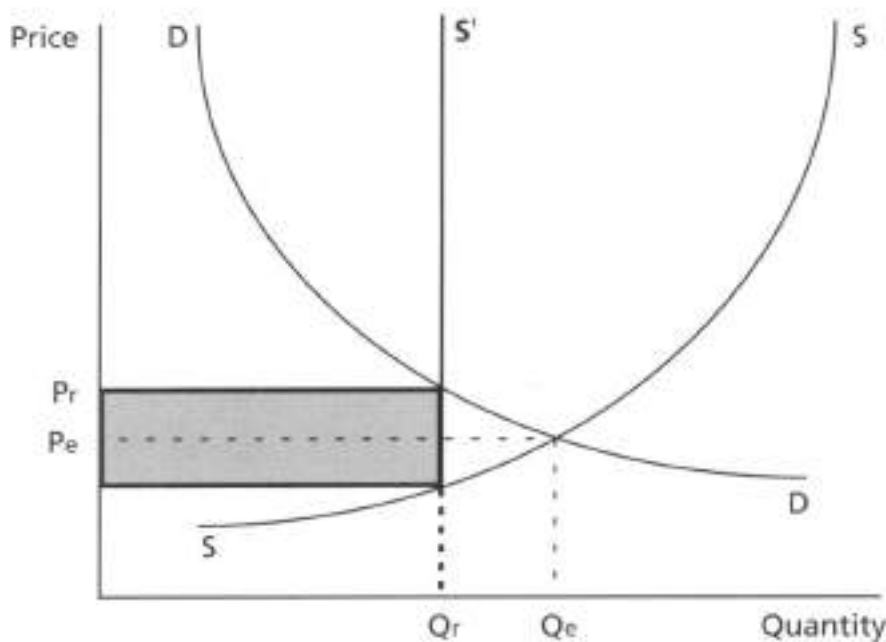
The second aspect of ease of entry relates to what is known as **indivisibility**. For some industries, the biggest impediment to entry is the size of the required investment. For instance, in the early part of the twentieth century, a profusion of car manufacturers opened for business. Because the technology was not that complex, many prospective producers entered the market. However, as technology advanced, and as Henry Ford radically altered manufacturing through his innovation of assembly line production, it became increasingly difficult for new manufacturers to spring up with new automobile designs; a huge investment was required to compete with the established car manufacturers. This impediment to market entry created an **oligopoly** (dominance by a few large sellers) in the automobile industry.

The outcome of the barriers to entry is known as **rent seeking**. When the market is restricted, any source of unusual profit attracts those who can exploit product scarcity and make a quick buck. Restrictions can creep in from a variety of sources, both public and private. In the public sector, restrictions on factor mobility result from government regulations. When there is a huge demand for a particular good but the government imposes a limit on its availability, the resulting excess demand causes its price to increase and revenues to climb. This increased revenue becomes excess profit for some. For those regulating the supply of the product, it can take the form of a bribe.

For those in the private sector, it provides a select group of business owners with the opportunity to make a monopoly profit. In the 1990s, California's milk producers successfully lobbied state politicians to set wholesale milk prices at an artificially high level and to impose restrictions on imports from neighboring states. In October 1999, the California Department of Food and Agriculture increased the wholesale price by 33 percent, one of the biggest increases in state history. The hike allowed California dairy farmers to earn an extra forty-two cents for every gallon of milk they sold to processors. As a result Californians spent about \$650 million more a year for milk than they would have in a free marketplace.¹³ In Figure 3.3, the shaded area represents this \$650 million rent.

Consider another example, drug trafficking in the United States. Without any government restrictions, the market price of an illicit substance would find its equilibrium, at which point the supply would satisfy all those who are willing to consume it at the going rate. However, when the government imposes sanctions, the supply is immediately restricted. This restriction increases its price and acts as a windfall to the drug smugglers—or as a bribe to corrupt law enforcement agents. In many developing countries, government corruption can be traced to officials' penchant for imposing too many restrictions on the economy (see "A Case in Point").

Rent seeking, therefore, leads to market inefficiency and can take many forms. Suppose your state awards a contract to a firm to build highways, at a profit deemed higher than normal by the firm's competitors. Say the normal profit for a similar project is \$10 million, but firm A will receive \$15 million. This situation may prompt firm A's competitor, firm B, to allocate up to \$5 million to fight the decision in court. If the decision is reversed in favor of firm B, and the firm spends only \$3 million in legal costs, then it makes a windfall

FIGURE 3.3 Rent Seeking

Note: The diagram shows the effects of an artificial restriction on supply. The curves DD and SS show the demand and supply of a commodity that, without any restriction, would reach its point of market-clearing equilibrium at price P_e and quantity Q_e . However, because the government has restricted the supply to SS , this action creates an artificial quantity decrease to Q_r and a price increase to P_r . The shaded area shows the possibility for rent seeking or corruption.

of \$2 million over what it would have earned in a competitive market. In the meantime, the project is delayed, and waste and inefficiency are introduced into the market.

Restricted or Asymmetrical Flow of Information

We are told that the secret formula for Coca-Cola is kept in a vault in Atlanta. In a perfectly competitive market, it would be anathema to have such a secret. However, contrary to the ideal constructs of a perfectly competitive market, a producer can often benefit from secrets. If the recipe for the fizzy drink had not been a secret and had been as available as the formula for hot dogs, another company might have posed stiff competition when the drink was introduced. But the secret kept competition at bay. In its early days, as the drink became wildly popular, other firms could not break into its newly created niche market. Later, with clever advertising and controlled distribution, the company established its influence over the soft drink industry.

Theories of perfect competition assume that information is free-flowing between buyers and sellers, so if you are in the market for a certain product, you can determine its true value to you. Yet in actuality, information is neither readily available nor free. In the case of some products, you can easily evaluate whether they are worth their prices. For instance, if you are considering buying a painting, you can look at it and determine if you like it or not. This type of good is called a **search good**—its quality is readily apparent. However, if the seller claims that it is an authentic painting by a reputed artist, and if you are not an art expert yourself, you may want to go to a certified appraiser and get it authenticated. To do so will cost you money—this is the cost of information. In the



A CASE IN POINT

Drug Policy: Our Addiction to Failure?

Since President Richard Nixon declared war on drug abuse in 1971, the nation has been spending nearly \$10 billion a year to combat the movement of illicit drugs from Latin America to the United States. The war on drugs has not only cost money and lives in the United States, it has also destabilized a large part of Latin America. The insatiable demand for drugs created the twin-headed monster of money and guns, which has devoured many to the south of us and is ultimately threatening U.S. national security. Money has corrupted the political systems of many Latin American countries, and lax gun laws in the United States have led to a steady stream of the most sophisticated guns flowing into Latin America, overwhelming the firepower of traditional police forces.¹⁴ At the time of writing, law and order are deteriorating rapidly in Mexico. Pictures of gruesome murders, including beheadings, committed by warring factions are becoming commonplace. In 2009 the U.S. State Department issued an advisory for Americans visiting Mexico, and many U.S. colleges discouraged students from crossing the border for traditional spring break revelry.

The profits from the drug trade have been the lifeblood of many violent organizations, such as the Revolutionary Armed Forces of Colombia (FARC), the Shining Path in Peru, and even the Taliban in Afghanistan, which directly threaten the United States. Although Colombia has seen a recent decline in violence and a weakening of FARC, and the Shining Path has been defeated in Peru, there is more cultivation of coca in Latin America and opium in Afghanistan.

In 2009, the Latin American Commission on Drugs and Democracy declared that despite trillions of dollars in spending, thousands of deaths, and hundreds of thousands of imprisonments, the war on drugs was "lost."¹⁵ It plainly pointed out that "prohibitionist policies based on eradication, interdiction and criminalization of consumption simply have not worked."¹⁶ In light of this failure, the commission proposes a paradigm shift based on three strategies:

1. Drug abusers should be treated as patients in need of public health services, rather than as criminals who need to be locked away.
2. The demand for drugs should be reduced by spending more on educating the public on the harmful effects of drug abuse.
3. The drug distribution network, run by organized crime, should be targeted more aggressively through police and intelligence work.

Discussion Points

1. What kinds of policies should we consider to combat illegal drug use? Public education about harmful effects of drugs? Amending gun laws to prevent guns from crossing borders and landing in drug traffickers' hands? Legalizing certain types of drugs?
2. What kind of metrics should we use to determine the success or failure of our policies?
3. What are the political, cultural, and historical impediments to a radical change in our drug policies?

eighteenth-century world of Adam Smith, such considerations were not as important as they are today. Therefore, in today's highly litigious society, the traditional warning of "buyer beware" is not sufficient, as it can lead to severe market inefficiencies.

In contrast to search goods, there are certain products you need to experience (or consume) before you can know their characteristics. By the time you have consumed a product and realized that it was not up to its billing and that the seller had withheld information about its limitations, it may be too late for you to receive compensation for your loss. These are called **experience goods**. We often hear about pharmaceutical companies being sued for selling drugs with unacceptable side effects that were not known to their users. Any drug will have some undesirable side effects. However, if a company knowingly engages in false advertising, which a consumer cannot be expected to verify, the government must step in and try to eliminate the market inefficiency.

Government attempts to correct these inefficiencies by passing proper disclosure laws, which can cover many aspects of trade, large and small. For instance, in several states, such as California, New York, and Missouri, a seller of real property must disclose any known occurrences of paranormal activities to a potential buyer. If you purchase a house that you believe is "haunted" and the seller was aware of this reputation at the time of sale, you can sue the seller for not disclosing the unwanted resident of your house.¹⁷ In another example, in 2009, the Food and Drug Administration (FDA) demanded that General Mills cease making its highly exaggerated claims of the cholesterol-reducing properties of its Cheerios cereal. The company claimed that its product could reduce "bad" cholesterol levels by 4 percent in just six weeks and ward off heart disease and cancers of the colon and stomach. These extraordinary claims were printed boldly on Cheerios packages, featured on the General Mills Web site and in television commercials, and even announced over loudspeakers in supermarkets as part of the company's promotional efforts. The FDA finally took the necessary step to stop the food giant from deceiving the consumers with its false claims.¹⁸

Externalities and Social Costs

Markets are efficient at producing goods that consumers want at prices they are willing to pay. This efficient process of allocating resources works well as long as an exchange between a buyer and a seller affects only those two parties. Unfortunately, their transaction often affects other parties. When it does, it creates an externality. The costs or benefits of individual externalities add up to what are known as **social costs** or **social benefits**. Externalities occur when the decisions of an economic actor (an individual, a firm, or a government unit) have an impact on another actor's well-being without the affected party's consent or compensation. If the impact is favorable, it is called a **positive externality**. If it is adverse, it is called a **negative externality**.

Suppose your town wants to attract a large retail outlet. The outlet is interested but would like a significant tax break. How can you justify using public money to help a privately owned company? Supporters of the proposal will point out the positive externalities: the creation of new jobs, the ability to attract other small businesses, and the influx of additional customers for bars and restaurants. These businesses will need employees, so even more jobs will be created. All of these activities will generate tax revenues for your town. They are examples of social benefits. Opponents of the project, on the other hand, will note the negative externalities, such as traffic congestion, an increase



A CASE IN POINT

College Affordability: Should Government Intervene?

The cost of higher education in the United States is rising much faster than the rate of inflation and parents' ability to pay for it, according to a report from the National Center for Public Policy and Higher Education.¹⁹ The report found that college tuition and fees increased 439 percent from 1982 to 2007, while the median family income rose 147 percent. As a result, college loans have more than doubled in the past decade. Compounding the problem, students from lower income families, on average, are getting smaller grants from the colleges they attend than students from more affluent families. While in many parts of the world, a child's future is tied to his or her parents' socioeconomic status, this country was founded on the principle of providing everyone the opportunity to prosper. Yet the rapid increases in college costs are threatening this "American dream" and some of our most cherished national goals.

We know that the completion of each successive level of education makes it possible for students to earn more money. In 2009, the U.S. Department of Education reported that while the average yearly income of a four-year college graduate was \$65,000, those with two-year college educations could expect to earn substantially less, \$38,000. This average goes down even further, as those with only high school diplomas earn, on average, \$26,000, and those without diplomas must get by on a meager \$18,000 a year.²⁰ It is no wonder that young men and women have looked at higher education as the key to success. This becomes particularly true during periods of economic downturn, when those who get thrown out of work attempt to increase their future earning capabilities by enrolling in institutions of higher learning, particularly state colleges and universities. Ironically, times of huge budget shortfalls cause state governments to cut back their support for higher education. In the spring of 2009, in the midst of the deepest recession since the days of the Great Depression, the University of California system announced a 9.6 percent tuition increase, while the California State University system also announced a 10 percent increase. Florida was talking about a 15 percent increase, while Washington State considered a

in crime, and the possible degradation of both the environment and the overall quality of life. These are the social costs of the retail outlet locating in your town.

The actor who generates positive externalities cannot redeem any of them. Similarly, those hit by negative externalities cannot recoup their losses through the market process. The private, or market, cost reflects the price of a good, the price we pay in stores. However, another cost may be associated with the consumption of this good. When I buy a pack of cigarettes, I pay the private, or market, cost. However, cigarette smoking is associated with a host of ailments that cost states huge amounts of money to care for patients suffering from these illnesses. When society (or the government, as the agent of society) pays for their care, that cost should be added to the product's market price. When it is not, the cost of the negative externality of smoking causes the social price to exceed the market price. That is, if you are a smoker, the price you are paying for the cigarettes is lower than what you should be paying when you take into account all the costs that you are imposing on the others. Since the market is unable to take corrective action on its

whopping 20 percent jump. The National Center for Public Policy and Higher Education reported that in 2007, the net cost of a year at a four-year public university amounted to 28 percent of the median family income, while a four-year private university cost 76 percent of the median family income. The situation is particularly daunting for those families with incomes in the lowest 20 percent; for them, the net cost of a year at a public university was 55 percent of the median income, up from 39 percent in 1999–2000. At community colleges, long seen as a safety net, that cost was 49 percent of the poorest families' median income in 2007–2008, up from 40 percent in 1999–2000. It is no wonder that in 2008 the cohort of twenty-five to thirty-four-year-olds was less educated than that of older workers.

According to a survey conducted by the Chronicle of Higher Education in 2008, 42 percent of Americans think that controlling college costs is "extremely important" and needs to be addressed by the federal government. Support for federal intervention in the rise of college tuition was strongest among parents (55 percent) and those with incomes under \$35,000 (65 percent).²¹ While lawmakers continue to debate intervention measures, such as creating a public watch list of universities whose tuition increases outpace inflation, the middle class has begun to assume higher debts for education costs, and many low-income students are left unable to afford college. As you can see, this phenomenal cost increase can tear apart some of the basic expectations of the people and create a permanent underclass, which, by any measure, is a harbinger of conflict and social discord.

Discussion Points

1. Why are college costs, particularly at public universities, rising so rapidly? Are these examples of market or government failure?
2. Should the government be involved in the education business? What should state governments do during the times of economic recession?

own, the government steps in with a tax to set the market price equal to the social cost. In the same way, when the government imposes taxes on a polluting firm, it bridges the gap between the market price and the social cost of the firm's product. In either case, the tax may be shifted to customers. The tax will reduce the demand for the good, thereby lowering the cost of consuming a socially harmful product.

Rising Service Costs: The Appearance of Market Failure

Another curious matter makes the functioning of the free market somewhat complicated. The cost of providing services—from education to health care, from personal services, such as plumbing, to car repair—is increasing much faster than are general commodity prices in the economy.²² This phenomenon, although not directly related to market failure, calls for government intervention. As a society, we are bewildered by rising health care costs and appalled at the cost of higher education—especially at public universities (see "A Case in Point"). We wonder why the costs of government

service delivery, at all levels of government, keep increasing, far outstripping the inflation rate. Moreover, as prices rise, the quality of many of these services seems to fall.

Although market failure and government intervention are responsible for some price increases, the major causes of such increases are rooted in simple economics, not “price gouging,” moral hazard (see chapter 1), or other kinds of market failure. Price increases also have very little to do with inefficiency in the government sector, though comparisons of government with the private manufacturing sector may suggest otherwise. For example, due to dizzying advances in computer technology, the cost of information processing is falling at an exponential rate. Technology is moving so fast that by the time this book is published, my one-year-old computer will have become obsolete. As we take advantage of computer-assisted design and manufacturing, worker productivity continues to shoot up at an astronomical rate. However, most service delivery by the government and public health providers remains impervious to technological advances. Even in our technology-saturated society, a postal worker still must deliver mail one house at a time, garbage must be picked up at the curb in front of each house, doctors must spend time with individual patients, and teachers must observe class-size restrictions. As a result, cost escalation is typical of many labor-intensive services, such as restaurant businesses, retail stores, and automobile repair shops.

GOVERNMENT FAILURE

In many ways the 1960s were magical, particularly for the United States. An awestruck nation watched astronaut Neil Armstrong take a “giant leap for mankind” to become the first person to step on the moon. The space race was over, and the United States had beaten the Soviet Union to the moon. The United States was indisputably on top in terms of military might and advancement in science and technology. However, in the midst of this exuberance, the nation faced widespread, endemic problems, from poverty and racism to political assassinations and street riots. A bewildered nation asked, if we can send men to the moon, why can't we solve our earthly problems? President Lyndon Johnson unveiled his Great Society program as a way to solve some of these vexing social problems. In the area of public budgeting, Johnson introduced the Planning-Programming-Budgeting System (PPBS), a process of allocating money through analytical means (see chapter 2). In the area of planning, physical planning based on technology was in vogue. In the realm of public policy, “systems analysis” was used to dissect every social problem into neat analytical components.

Unfortunately, the government failed to deliver. The Great Society program did not eliminate poverty,²³ the PPBS was a dismal failure, physical planning that emphasized building new public housing and freeways was unable to cure urban blight, and systems analysis grappled unsuccessfully with social problems. It was indeed easier to remove technological hurdles to landing on the moon than to find solutions to the age-old problems of human society. If in an imperfect world the market often fails to deliver, so does the government. There are ten reasons for this failure, each of which we will discuss briefly:

1. Inability to define social welfare
2. The limits of democracy and the paradox of voting

3. Inability to define the marginal benefits and costs of public goods
4. Political constraints
5. Cultural constraints
6. Institutional constraints
7. Legal constraints
8. Knowledge constraints
9. Analytical constraints
10. The timing of policies

Inability to Define Social Welfare

Recall the lofty list of inalienable rights in the Declaration of Independence. The sentence beginning “We hold these truths to be self-evident” boldly assumed that “we” included everybody in the new nation. Yet history has shown repeatedly that not everybody in the nation has held the fundamental rights guaranteed by the Declaration, rights deemed universal regardless of race, religion, sex, or national origin.

Similarly, we frequently hear phrases such as “the welfare of society” or “the good of the people” in political rhetoric or social discourse. Yet in a secular, pluralistic political culture, a thoughtful reader of public policy analysis recognizes the fragility of such confident phrases. We can think of no policy that is a boon to every individual in society. Instead, we are apt to find cases in which there are winners and losers, victors and vanquished, benefits and costs (see chapter 14).

The Limits of Democracy and the Paradox of Voting

If we cannot define “social welfare,” then we can assume that people will articulate their wishes through the electoral process. Let us say that a rural county has recently experienced a spate of devastating forest fires. It is now trying to decide how many fire engines to buy—one, two, or three. The county population is divided over the issue. Three public opinion groups have emerged: those in a small minority who are most threatened by the possibility of another fire next summer (group A), those who are more concerned about the solvency of the county than about the prospect of another fire (group B), and those who are taking a centrist position (group C).

As you can see in Table 3.1, if a referendum were held and each group were allowed to vote for its first choice, we would observe no conclusive outcome. Although the moderate group is numerically the strongest, it does not command the majority of the votes. However, if each group were to vote on two alternatives, we would discern a conclusive majority preference. When the choice is between two fire engines or three, the two-engine option receives 80 percent of the votes (groups B and C prefer two to three). When groups must decide between one engine or two, two still wins (garnering 65 percent of the votes), because those who are most concerned about another forest fire would rather have two engines than one. In this case, a social consensus will be reached, and the outcome will reflect the will of the majority.

However, the situation gets more complicated if group A changes its preference ranking. Suppose group members are so frightened by the prospect of living without

TABLE 3.1 **Voters’ Preferences on the Decision to Purchase Fire Engines**

Group	Strength (percentage of votes)	Number of fire engines		
		First choice	Second choice	Third choice
Those most threatened by the prospect of fire (A)	20	3	2	1
Fiscal conservatives (B)	35	1	2	3
Moderates (C)	45	2	3	1

adequate fire protection that they vote for only one fire engine if they cannot have three. Perhaps in protest they move to another county.

Suppose that group A, the smallest and most intense, finds the compromise position of two fire engines just as unacceptable as that of only one (see Table 3.2). In that case, group members have made the strategic decision to rearrange their preference pattern. Now they do not prefer two to one, even though two is closer to their desired position of three. You can see that there will be no unique outcome of this election. The outcome will depend on who controls the agenda:

Choice between 1 and 2	1 wins with	55 percent of the votes
Choice between 3 and 1	3 wins with	65 percent of the votes
Outcome	3 wins	
Choice between 3 and 2	2 wins with	80 percent of the votes
Choice between 1 and 2	1 wins with	55 percent of the votes
Outcome	1 wins	

As a result of this new strategic position, group A has effectively eliminated the middle position. If it can control the agenda, it can win the election with only 20 percent of the vote! This is the paradox of voting. Its broad implications are obvious: *even with the fairest voting rules, the outcome of an election may not reflect the true preferences of the community.*

Inability to Define the Marginal Benefits and Costs of Public Goods

If we are at a loss defining “social welfare,” we are similarly confounded defining the marginal utility from an additional dose of public expenditure. When President Ronald Reagan took office, during the height of the Cold War with the Soviet Union, he declared that he wanted a 700-ship navy. Nearly a decade later, Bill Clinton, presiding over the lone superpower, opted for a navy of fewer than 400 ships, nearly half the number requested by Reagan. Why those numbers? Why not 701 ships or 420 ships? Would you, as a voter, sleep more soundly if you knew that one more navy destroyer than before

TABLE 3.2**Voters' Preferences on the Decision to Purchase Fire Engines: Strategy Change by Group A**

Group	Strength (percentage of votes)	Number of fire engines		
		<i>First choice</i>	<i>Second choice</i>	<i>Third choice</i>
Those most threatened by the prospect of fire (A)	20	3	1	2
Fiscal conservatives (B)	35	1	2	3
Moderates (C)	45	2	3	1

were guarding your shores? These questions reveal the difficulty of calculating the marginal utility of a government purchase.

Political Constraints

In his first meeting with President Franklin D. Roosevelt, Keynes did not seem practical to the savvy politician. In Keynes's mind the policy implications of his economic strategy—expand government spending, even beyond revenues, during times of recession and reduce expenditures and government size during periods of inflation—did not affect the soundness of his theory. But as Keynes, Roosevelt, and the rest of the nation soon learned, it is much easier to raise the levels of expenditure than to reduce them. Every time policymakers create a government program, the program creates its own constituents. As a result, even when its utility is in question, closing it down often becomes nearly impossible.

Cultural Constraints

Like political constraints, the cultural context influences the public policy process. To be adopted and successfully implemented, public policies must conform to the cultural norms of the community. This cultural context perplexes us when we travel across the nation or to another country. A visitor to India might wonder why cows are allowed to roam freely through the already overcrowded city streets. In the same way, outsiders in America might ask about the cultural attachment of ordinary Americans to guns in the face of gun-related violence. They might also question those who oppose international abortion and birth control initiatives, even as United Nations population control programs in poor, lesser developed countries buckle under unsustainable population growth. Without cultural acceptance, even a perfectly reasoned public policy may not be considered appropriate for a community.

Institutional Constraints

Any public policy must depend on bureaucratic institutions for its formulation and implementation. An organization, like any other entity—collective or individual—develops its own cultural ethos, goals, and mythology. As a result, policies promoted by the social services division of a city can come into direct

conflict with the mandates of its law enforcement branch. These conflicts, often seen as “turf battles,” can render a policy ineffective. The Commerce Department promotes international trade. Therefore, it seeks to maximize the export of American goods. Yet its goals may collide with those of the State Department, which manages the country’s foreign policy. For political reasons the State Department may bar U.S. companies from selling their wares, such as weapons and high-speed computers, to other countries.

Legal Constraints

Public policies must be formulated and implemented within a nation’s legal framework. Generally speaking, laws from six sources govern our daily lives in the United States: constitutional laws, laws made by legislatures, executive orders, interpretations of laws by the judicial branch, agency rules, and public referenda. In our democratic system of checks and balances, laws passed by legislatures, executive orders, referenda, and agency rules can all be declared null and void by the courts. In certain cases the legislative branch fights the executive branch for control of the national agenda. Although Congress passed the War Powers Resolution of 1973 to curtail the president’s ability to send U.S. troops into combat in other countries, it failed to enforce its own laws when President George W. Bush sent U.S. forces to fight in Iraq and Afghanistan.

Knowledge Constraints

Any public policy is constrained ultimately by our present knowledge. As we learn more about the world around us, we realize the failings of our past policies. Take the problems of malaria abatement in many countries around the world. Malaria was once the top killer in the world’s tropical regions. As scientists learned that the disease spread through mosquitoes, they sparked a huge international effort to kill the pesky insects in their places of breeding—swampy areas with stagnant water. Thanks to the massive spraying of the insecticide DDT, many lesser developed countries declared victory over malaria’s scourge. Yet within a decade people started noticing the chemical’s devastating effects on the fragile ecology of marshland, which is essential to the survival of many species of flora and fauna. Again, policies had to be developed to eradicate not the disease but the effects of past policy. As these devastated habitats were restored, the disease came back, often with a vengeance.

Analytical Constraints

Analytical techniques tapped for public policy analysis, either at the formulation stage or the implementation stage, use numbers. Numbers have a magical quality. They give the impression of being totally objective. As we will see, however, there are many opportunities for subjectivity to creep into our analysis. When we learn statistical methods or the various techniques of operations research, they exude the impartiality of scientific reasoning. Yet when we dig deeper, we discover that *objectivity relates solely to deriving the conclusion once the problem has been formed*. In each stage of formulating the problem, we may confront not just confusion but also pressure to make quick decisions, which are often rendered for the convenience of the analyst or to suit the particular

quantitative technique chosen for the analysis. These analytical constraints may creep into the analysis from a number of sources discussed throughout this book.

The Timing of Policies

Biases or inefficiencies may arise due to the timing of policy measures. Consider a business cycle that moves the economy through a series of booms and busts. Critics of government's anticyclical policies point out that several steps must be taken before a policy is implemented and its effects felt. If the government decides to enact countermeasures to keep the economy on an even keel, it must first recognize that the economy is going into a recession. As we note in chapter 4, though, forecasting is sometimes like looking in the rearview mirror while driving a car. We consult past data to gauge where we should be going. But the data have been collected from the quarter before. We encounter a **recognition gap**, which requires that we study trends to understand its causes. Because it takes time to analyze a trend, unless we are studying a rapidly escalating situation, policies can become delayed. A time gap often occurs as analysts come up with appropriate policies. This gap is called a **prescription lag**. After a policy has been recommended, the appropriate legal body must consider it for adoption, causing an **adoption lag**. Finally, an **implementation lag** results when adopted policies are translated into action. In the meantime, if the economy has made a turn on its own, the "lagged" policies may backfire and choke off a natural recovery. This policy failure then creates a new set of problems.

The timing of an anticipated change in government can also lead to market inefficiency. Thus, if I expect the market price of an item to go up, I may run to the store and stockpile, buying in quantities that I may not need for some time. Furthermore, if this fear is widespread, the surge in market demand can create shortages and an artificial increase in prices. For instance, immediately after the election of Barack Obama, many gun enthusiasts feared a looming restriction on gun and ammunition sales. This caused a frenzy in sales of guns and even created a shortage of ammunition for public safety officials. As a result, many even jokingly called Obama the "Gun Salesman of the Year."²⁴

LIMITING GOVERNMENT INTERVENTION: JOINT PARTNERSHIP IN A MIXED ECONOMY

The standard theory of economics tells us that the market achieves its highest level of efficiency when government resists the temptation to interfere in it. Yet this theory yields many questions. Why does the market go through cycles of boom and bust? Why does unemployment exist even when there is no minimum wage to discourage firms from hiring workers? Why does an unchecked market produce extremes of wealth and poverty? Rather than attributing these problems to government interference, Keynes thought they were caused by humans' tendency to respond "irrationally" to a number of cues. He called this the **animal spirit**. Building on this, the Nobel Prize-winning economist George Akerlof and the Yale economist Robert Shiller argue that if the market is to avoid the disruptive influences of the animal spirit, it must understand human nature in its fuller context and not be tied to its restrictive assumption of self-utility maximizing *homo economicus*.²⁵

However, government interference in the market is a double-edged sword, and sometimes governments do fail. When this happens, because of government inefficiency or corruption, the market forces of unencumbered demand and supply must be brought in to remedy the situation. The reintroduction of market forces can either stimulate or simulate the market. **Market stimulation** takes place when the barriers of entry into the market are removed, whereas **market simulation** requires the government to provide public goods through exclusionary pricing. Akerlof and Shiller insist that the ill effects of the animal spirit can be brought under control only through a properly balanced partnership between the market and the government.

Stimulating the Market

Market stimulation is achieved when the government *deregulates, legalizes, privatizes, or infuses money*. As we have seen, government intervenes in the market through regulations when it deems such actions to be beneficial to society. One of the most regulated sectors in U.S. economic history is the banking industry. Through their lending, commercial banks create an economy's money supply. As the backbone of the monetary structure, banks can function only when people trust the system and deposit money. When we open a checking account, the Federal Deposit Insurance Corporation requires the bank to pay a small premium to buy insurance against a possible bank failure. If the bank fails, each of the checking accounts, up to a maximum of \$250,000, will be fully insured. When bank insurance was originally proposed during the Great Depression, President Roosevelt was deeply suspicious of it. He feared that because customers could not lose a penny, bankers would be irresponsible with deposited money. In the 1980s, the so-called supply-side economists successfully advocated **deregulation** of the market, which erased many of the restrictions that had prevented banks from engaging in highly speculative activities, such as selling risky commercial stocks or investing in commercial real estate. A decade after deregulation, the savings and loan crisis engulfed the nation. Many critics blamed the freewheeling practices of unregulated banks and sought to reregulate the sector.²⁶ As we have seen in the previous discussion of the failure of the insurance giant AIG (see "A Case in Point" earlier in this chapter), many of these same mistakes have been repeated in recent years.

Another area that is heavily regulated by the government is electricity. Regulations have caused inefficiency, prompting the U.S. government to deregulate the industry. Although advocates of deregulation often tout its virtues, evidence indicates that it has had mixed results. In North Carolina, the deregulation of electricity did not reduce consumers' rates.²⁷ In fact, in the late 1990s, largely because of deregulation, electricity rates in fifty-one cities jumped by 25 to 30 percent. In winter 2001 the deregulation of electricity brought California perilously close to an economic disaster. Not surprisingly, one study found that deregulation of the airline industry failed to increase efficiency as had been predicted.²⁸ It is extremely difficult to establish unequivocally a causal relationship between deregulation and market failure, so the debate on deregulation's merits is likely to go on.

Legalization brings a previously illegal act into the realm of taxable, legal activities. The Supreme Court's judgment in *Roe v. Wade* legalized abortion for American women within a certain period of pregnancy. This action drastically reduced the number of

unsafe abortions. However, as the opposition to the Court's opinion in *Roe v. Wade* demonstrated, the acceptance of a public policy may not always be based strictly on economic grounds or the health considerations of a pregnant woman but can come from a host of moral and religious perspectives. Similarly, in certain European nations, as well as some counties in Nevada, laws permit prostitution. This generates tax revenues for these governments (as the sex workers work within the legal system), lowers the risk for spreading sexually transmittable diseases, and may restrict the power of pimps and organized crime. Yet many in the society recoil at the idea on moral, ethical, and religious grounds.

Privatization turns previously government-operated services over to the private sector. The conservatism prevailing in the nation following the election of President Reagan spurred a huge demand for the privatization of government services. Although the initial enthusiasm waned with time, privatization continued to advance at all levels of government.²⁹ Despite the common assumption that privatization is a quick way out for cash-strapped governments, it also has been used by solvent governments to acquire superior management talent and specialized expertise. However, despite the growing interest in privatization, fierce political opposition, particularly from labor unions, continues to make outsourcing a hard sell. Proponents of privatization say that competition from private firms can often raise productivity in government-provided service areas. In the municipal waste disposal industry, evidence suggests that parallel providers have helped both private and public sector firms become "better, faster, smarter, and more cost-effective." Residents of cities and counties across the United States have been the primary beneficiaries of this competition.³⁰

Government also tries to stimulate the market through the **direct infusion of money**. In times of recession the aggregate demand for goods and services goes down. As a result, market prices fall, and industries start laying off workers. As these unemployed workers lose their purchasing power, prices go down even more, prompting another round of layoffs. This process continues, sending the entire economy into a downward spiral. Keynes pointed out that, unlike the situation shown in Figure 3.1, there is no market mechanism that can bring the economy out of this whirlpool of an economic abyss. The only way the economy can be saved is by the government's borrowing huge sums of money and injecting them into the economy. As the government funds projects, which employ the previously jobless, the reverse process takes hold: workers spend their money and boost aggregate demand, which increases prices and profits businesses. The Obama administration's decision to pass an enormous stimulus package is rooted in Keynesian theory. On the other hand, those who believe in the market argue that government intervention only brings about inefficiency, corruption, and public debt, which binds future generations to a higher level of taxes.

Simulating the Market

When it is not possible to stimulate a market, federal, state, and local governments reduce the inefficiencies of government control by simulating the market—that is, by charging prices for the products and services once offered free of charge (see "A Case in Point"). Charging fees for government services has become critical for cash-starved cities in many parts of the country. This has been particularly true in California, where in

1978 voters organized a "tax revolt" and passed a sweeping referendum, known as Proposition 13, that reduced their property taxes. This proposition also significantly reduced the government's ability to impose new taxes. Because the voters said no to the government's request for finances but did not at the same time curtail their demand for services, the only answer to the problem was to charge user fees. More recently, some cities, such as Coon Rapids, Minnesota, have considered charging their citizens when they use emergency services like fire extinguishing or cleanup after car accidents.³¹ With government as the service provider, users pay for the services they demand, simulating the market process.

Apart from allowing financially strapped cities to provide the services their citizens demand, simulating the market through user fees offers several other advantages. First, prices act as a signal for demand. Because it is impossible to quantify the demand for public goods, if a local government charges a fee for its services, each consumer can bring his or her demand for public goods in line with their cost. Returning to our example of voting for fire engines, we may satisfy all three groups by charging them different rates for different levels of fire protection. The group that wants the most fire protection can pay for extra fire engines dedicated to their community, while the rest are charged the regular rate. This situation is known as **consumption efficiency**, in which an economic actor's marginal utility matches the marginal cost. Because nothing dispels enthusiasm more than a small entry fee, user fees often can curb the abuse of a publicly provided service.³²

Along with consumption efficiency, the simulation of the market may also introduce **production efficiency** by forcing government units to bring their marginal revenues in line with marginal costs. In fact, sending price signals through user fees may help even the private market correct itself. In the market the social cost of a product often becomes significantly higher than its market cost. Suppose the authorities of the indoor sports arena in your town want to host a pop music festival lasting several days. Although it may be profitable, this concert may impose extra costs on the surrounding areas in terms of higher street-cleaning expenses and more traffic, pollution, and crime. Each business may have to hire extra help to clean up or protect its property during the festival. In addition, the city may have to pay police officers overtime to direct traffic and keep the festival orderly. As you might expect, festival organizers pass on their costs to the rest of the community. Therefore, if the city charges the organizers a fee for the privilege of holding such an event, an amount that compensates the affected parties, no one will be economically harmed, yet many (promoters, stadium authorities, musicians, and fans) will benefit from the festival. This strategy is effective for many nonprofit organizations, such as charitable, religious, and educational institutions, which are exempt from property taxes. By imposing user fees, they can bring their private costs in line with those of society.

The decision to charge for publicly provided services must be considered very carefully. Some services provide benefits not only to those who use them but also to the rest of the community. As a society we have decided that when children in grades kindergarten to 12 are educated, the benefits are not restricted to them but spill over to the rest of society. Therefore, it is not always prudent to emulate the market and charge tuition fees. Charging a fee may have an adverse effect on poor people. If a city wants to impose a fee for using public facilities such as basketball courts and baseball fields, the fee may



A CASE IN POINT

Trading Pollution: Simulating the Market

Ronald Coase of the University of Chicago, who won the Nobel Prize in economics in 1991, had a novel idea. He argued that if people were allowed to market their property rights, many problems resulting from negative externalities could be solved.³³ Take the case of environmental pollution.³⁴ Factories emit pollutants—sources of negative externalities—as by-products of their operations. However, some factories are more efficient than others and produce fewer pollutants. Regulating environmental pollution has a high opportunity cost because it causes production shutdowns, loss of income, and unemployment. If there are too many restrictions, the government imposes undue burdens on society, but if the laws are too lax, the social costs of the products far outweigh their market costs.

Guided by Coase's arguments, government experts determine the amount of pollution that can safely be tolerated, then give each polluting plant a permit to pollute up to its quota. If some plants pollute below their quota, they can sell their "rights to pollute" to plants that cannot meet the regulatory standard. Through trading pollution rights, the social costs can be brought in line with the market costs. Thus, the Clean Air Act of 1990 introduced what is now known as emissions trading, by which the government sets pollution quotas and efficient firms sell their unused pollution rights to inefficient ones.

In response to pressure to reduce the emission of greenhouse gases, BP-Amoco's Western Gas business unit installed thousands of new valves that shunt fluids from the firm's network of wells. The old valves emitted whiffs of methane, a potent greenhouse gas; the new valves do not. Today, Western Gas sells its unused pollution rights to other subsidiaries of Amoco at a profit.³⁵ This sort of emissions trading has been steadily increasing every year. In 2005, 374 million metric tons of carbon dioxide equivalent were exchanged, a 240 percent increase over the 2004 level.³⁶ Although by many measures this injection of market forces into the allocation of public goods has been a success, critics have argued that the overall cap on total emissions is too high. As a result, some firms have found new ways of making money by spreading pollution in previously unpolluted parts of the world.

Discussion Points

1. How can we bring social costs close to market costs through government action?
2. Do you think trading emissions offers a viable solution to the problem of global warming?
3. Can you think of other negative externalities that could be treated with Coase's prescription?

have an adverse effect on the children of a poor neighborhood. In contrast, if the city wants to raise the fee for using public golf courses, such effects may not be felt.

Even in the face of overuse or high social cost, charging a fee may not be technically feasible, or the cost of collection may be too high. Suppose your city's traffic department

wants to reduce congestion during peak traffic hours. It may make sense to charge motorists a fee for using the highways during those hours.³⁷ However, it may not be possible to collect those fees without creating an even bigger traffic jam. In such cases, simulating the market is not a viable option.

Besides being technically unworkable, imposing fees for services that previously were provided free of charge may not be politically feasible. If services are free, consumers view them as having no cost. Toward the end of the nineteenth century, San Diego faced a small tax revolt of its own. The city, like most other cities in the nation, was collecting fees for picking up garbage. When it became known that the city was selling this garbage to nearby hog farms, irate citizens passed a referendum to eliminate garbage collection fees. More than 100 years later, the city has grown substantially and the pig farms have long since left the vicinity. Yet it remains politically infeasible to reinstate a collection charge, one of the most common municipal fees in the United States.

KEY WORDS

Adoption lag (p. 67)	Market stimulation (p. 68)
Animal spirit (p. 67)	Mercantilism (p. 47)
Consumer surplus (p. 50)	Negative externality (p. 59)
Consumption efficiency (p. 70)	Oligopoly (p. 56)
Deregulation (p. 68)	Opportunity costs (p. 49)
Direct infusion of money (p. 69)	Positive externality (p. 59)
Equilibrium (p. 48)	Prescription lag (p. 67)
Experience goods (p. 59)	Price maker (p. 53)
Externality (p. 53)	Price taker (p. 50)
Factor mobility (p. 55)	Privatization (p. 69)
Implementation lag (p. 67)	Production efficiency (p. 70)
Indivisibility (p. 56)	Recognition gap (p. 67)
Legalization (p. 68)	Rent seeking (p. 56)
Marginal analysis (p. 49)	Search goods (p. 59)
Marginal cost (p. 50)	Social benefits (p. 59)
Marginal revenue (p. 50)	Social costs (p. 59)
Market simulation (p. 68)	X-inefficiency (p. 54)

EXERCISES

1. Your city wants to contract out fire prevention and firefighting services to a private firm. Argue for and against such a move.
2. Write an essay on the uses and limits of economic analysis, incorporating an appropriate example from your local or state government.
3. Write an essay on the deregulation of utility prices. When do you think deregulation works, and when does it not?

Notes

1. Stephanie Ebbert, "Menino Bill Would Boost Low-income Housing," *Boston Globe*, November 30, 2000.
2. The instant success of the book—it went through five editions during Smith's lifetime—made the author famous, but his work failed to have a significant influence on British public policy of the time. For a discussion of Adam Smith's contribution to modern economics, see David D. Raphael, *Adam Smith* (Oxford, United Kingdom: Oxford University Press, 1985).
3. Smith did not draw the famous demand and supply curves. Developing the geometry of his theory took about 100 more years and a combined effort on both sides of the English Channel. For a detailed discussion, see Mark Blaug, *Economic Theory in Retrospect* (Homewood, Ill.: Richard D. Irwin, 1968), 38–67.
4. Ebbert, "Menino Bill."
5. Daniel Yergin, *The Prize: The Epic Quest for Oil, Money, and Power* (New York: Simon & Schuster, 1991).
6. See Liam Plevin, "AIG Sees Long Road Back from the Brink," *Wall Street Journal*, May 11, 2009.
7. See the World Bank's "Key Development Data & Statistics" page at www.worldbank.org/data/ (select "Data" from the menu at the left side of the page).
8. Stephen Labaton, "Administration Plans to Strengthen Antitrust Rules," *New York Times*, May 11, 2009.
9. Joel Brinkley, "The Verdict: Microsoft Is a Monopoly," *New York Times*, November 7, 1999.
10. Harvey Leibenstein, *Beyond Economic Man: A New Foundation for Microeconomics* (Cambridge, Mass.: Harvard University Press, 1976). See also Roger S. Franz, *X-Efficiency: Theory, Evidence and Applications* (Boston: Kluwer Academic, 1988).
11. Kate Gillespie and Clement Moore Henry, eds., *Oil in the New World Order* (Gainesville: University Press of Florida, 1995).
12. There is no single, worldwide definition of unemployment. The official U.S. definition carries many restrictions, and the definitions of other nations vary significantly.
13. The Department of Food and Agriculture defended its action by arguing that it was only trying to ensure the survival of California's dairy farmers. Although this rhetoric evokes the image of small family farmers getting up with the sunrise every morning to milk their herds, in fact, dairy farming is a huge, \$3.6 billion industry. The average farmer is a corporation with 602 cows and annual sales of \$1.6 million.
14. See Brian Loveman, ed., *Addicted to Failure: U.S. Security Policy in Latin America and the Andean Region* (Lanham, Mass.: Rowman & Littlefield, 2006). Also see David R. Mares, *Drug Wars and Coffeehouses: The Political Economy of the International Drug Trade* (Washington, D.C.: CQ Press, 2005).
15. Drug Policy Alliance Network, "The Latin-American Commission on Drugs and Democracy (Co-chaired by former presidents Fernando Henrique Cardoso [Brazil], César Gaviria [Colombia], and Ernesto Zedillo [Mexico]) Releases Groundbreaking Report," February 11, 2009, www.drugpolicy.org/news/pressroom/pressrelease/pr021109lar.cfm. For an interesting demonstration of the costs of the drug war, see Drug Sense, "Drug War Clock," www.drugsense.org/wodclock.htm.
16. Fernando Henrique Cardoso, César Gaviria, and Ernesto Zedillo, "The War on Drugs Is a Failure," *Wall Street Journal*, February 23, 2009.
17. In a famous case, often called the *Ghostbusters* case, a seller in New York was held liable for withholding information about the existence of poltergeists from the buyer (*Stambovsky v. Ackley*, 169 A.D.2d 254, N.Y. App. Div. 1991). This case has become a textbook example in contract and property law cases. There have since been other cases of similar litigation.
18. A large literature on information asymmetry and the cost of information began with the publication of George Akerlof's "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics* 84 (1970): 488–500.

19. John Immerwahr and Jean Johnson (with Paul Gasbarra, Amber Ott, and Jonathan Rochkind), "Squeeze Play 2009: The Public's Views on College Costs Today," Public Agenda and the National Center for Public Policy and Higher Education, 2009, www.highereducation.org/reports/squeeze_play_09/index.shtml.
20. U.S. Department of Education, www.ed.gov/index.jhtml.
21. Jeffrey Selinger, "Americans Split on Government Control of Tuition," *Chronicle of Higher Education*, April 4, 2008.
22. This concept of productivity lag was first presented by William J. Baumol, "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *American Economic Review* (June 1967): 415–426.
23. Although the poverty rate did decline during periods of economic expansion, poverty was merely cut in half between 1959 and 1972.
24. Judson Berger, "Obama Driving Surge in Gun Sales, Firearms Groups Say," FoxNews.com, January 16, 2009, www.foxnews.com/politics/2009/01/16/firearms-associations-claim-obama-drove-surge-gun-sales.
25. George A. Akerlof and Robert J. Shiller, *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism* (Princeton, N.J.: Princeton University Press, 2008).
26. See, for example, Alan Gatt, *Regulation, Deregulation, Regulation: The Future of the Banking, Insurance, and Securities Industries* (New York: Wiley, 1994).
27. Buster Kantrow, "Electricity Deregulation Runs into Static in N.C.," *Wall Street Journal*, July 26, 2000.
28. Zinan Liu and E. L. Lynk, "Evidence on Market Structure of the Deregulated US Airline Industry," *Applied Economics* 31 (1999): 1083–1092.
29. "Privatization on the Rise Despite Surpluses," *USA Today*, August 1, 1999.
30. Christi Clark, "How Privatization Helped Raise the Bar in the Solid Waste Field," *American City and County* 113 (February 1998): 52–60.
31. Maria Elena Baca, "Cities May Charge Fees on Emergency Services," *Minneapolis Star-Tribune*, May 2, 2009.
32. Although this statement is made somewhat in jest, there is strong evidence that people are willing to pay for a public resource if doing so improves service. Kelly McCollum found in a survey of 3,000 students in the California State University system that about 65 percent of the students would be willing to pay higher than usual student fees for better on-campus computing resources. Of those students who were willing to pay, more than 71 percent stated that they would spend an extra \$10 per month for better on-campus computing, while 15 percent were ready to pay more than \$20 per month. See Kelly McCollum, "65% of California State U. Students Back a Technology Fee, Survey Finds," *Chronicle of Higher Education*, September 17, 1999.
33. For an excellent discussion of the theoretical issues, see Neil Bruce, *Public Finance and the American Economy*, 2nd ed. (Boston: Addison-Wesley, 2001), 99–115.
34. A. Denny Ellerman, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth M. Bailey, *Markets for Clean Air: The U.S. Acid Rain Program* (Cambridge, United Kingdom: Cambridge University Press, 2000).
35. John Carey, "A Free-Market Cure for Global Warming," *Business Week*, May 15, 2000.
36. See the World Bank's Carbon Finance Unit (www.carbonfinance.org) for the latest data and information.
37. In some parts of the country transportation departments have experimented with issuing permits for single motorists to use carpool or high-occupancy vehicle (HOV) lanes. These permits are often expensive because they must pay for highway patrol officers to enforce HOV restrictions.

4

The Policy Process

Something interesting is happening on the street corners of my city, though I did not notice it at first while driving through my neighborhood. I approached a yellow light, and instead of coming to a stop at the signal, I made a quick turn a split second after the light had turned red. Three weeks later, I received an envelope from the Department of Motor Vehicles. Inside were a picture of my license plate and a clear shot of this author looking intense while making an illegal turn in a hurry. The pictures came with a hefty fine and an invitation to spend a weekend at traffic school. On my next trip to that same intersection, I noticed the little birdhouse-type structure at the corner, its shiny camera lens sticking out, busily snapping away at delinquents. Within a few months I observed that most of the city's intersections, and many in neighboring towns, had been outfitted with these Orwellian devices. In June 2000 the cameras operated in forty-one U.S. cities; today, nearly all major U.S. cities use them.¹

These devices clearly were installed as a result of deliberate public policies. Students of policy analysis might ask, who placed the cameras on the policy agenda? How did policymakers decide to install them? How are city officials implementing the policies? How effective are the devices at preventing traffic violations? As the cameras become part of our daily lives, we may ask whether they raise constitutional issues. If so, how might this concern change future policies for catching red light runners? Finally, if policymakers find the cameras ineffective, how do they go about eliminating them? We may answer all of these questions by examining the **policy cycle**, or **policy formation**, the process through which policy is formed (see Figure 4.1).²

Before examining each stage of the cycle, we must remember an important caveat: the process is not as neatly segmented as the figure suggests. Rather, it is much more complex, resembling a seamless quilt in which all phases of the cycle are woven together.

FIGURE 4.1 The Policy Cycle

AGENDA SETTING

The policy cycle starts when the government pays serious attention to an issue. If you look around, you notice that thousands of areas cry out for corrective action. However, not everything that needs the government's attention gets it. You may notice a particularly dangerous intersection in your neighborhood. When you turn into the intersection, your view is blocked by a large tree. You are concerned, and so are some of your neighbors. Yet years go by and no one takes action. Or you may worry about the so-called greenhouse gases that reportedly cause global warming. You may wonder why political leaders do not pay more attention to a problem with such enormous consequences for our future.³ You may also ask why some issues suddenly crop up, dominate the agenda, and surprise everyone, including those who study such matters for a living.⁴

Policy scientists identify two kinds of agendas: those that government institutions act on, known as institutional, or governmental, agendas, and those on which they delay action, called systemic, or noninstitutional, agendas.⁵ **Systemic agendas** "percolate" in society, waiting to be elevated to the active agenda. A systemic agenda "consists of all issues that are commonly perceived by members of the political community as meriting public attention and involving matters within the legitimate jurisdiction of existing governmental authorities."⁶ In contrast, an **institutional agenda** consists of those problems on which "legislators or public officials feel obliged to take appropriate measures."⁷ Sometimes an issue remains the focus of intense public debate but still never makes it onto the institutional agenda. At other times an issue surprisingly and suddenly moves to the active list. Take, for example, the issue of teen smoking. Despite the fact that for years many prominent health care professionals had documented tobacco's ill effects and highly addictive qualities, the issue failed to reach the institutional agenda until the late 1990s, when some states sued tobacco companies, the Food and Drug Administration issued stronger tobacco regulations, and the media covered the issue more widely—Congress responded by passing measures that sought to reduce underage tobacco use by prohibiting certain types of cigarette advertisements and reforming retail distribution practices.

The White House and Congress initiate official or institutional agendas at the federal level. The courts also set institutional agendas through their decisions and legal opinions. Because of the system of checks and balances, all three branches of government must cooperate to adopt an institutional agenda. All legislation must be passed by both houses of Congress. If a bill conflicts with the Constitution, the judiciary must step in. Just as government actors often pull institutional agendas from the vast pool of systemic agendas, they may at any time toss back an institutional agenda into the pool.

How does an issue reach the active agenda, and who ushers the issue through the policy cycle? Those who hold demand-side theories of agenda setting believe that policymakers determine institutional agendas in response to widespread popular demand. This view is known as the **pluralist model**. In contrast, those who support the **elitist model**, a supply-side view, believe that the power to set the agenda resides at the top of the political and economic hierarchies.

The Pluralist Model

The pluralist model is based on theories of American democracy developed in the 1950s and 1960s.⁸ In this somewhat romanticized view of the American power structure, power rests with citizen activist groups. The model makes no permanent distinction between the elite and the masses, though a hierarchy may exist in which the elite control some institutional agendas. Any issue can reach the institutional agenda, and its success in doing so depends on the abilities of citizen activist groups (see Figure 4.2).

In the early 1970s Roger Cobb and his associates explained how an inactive agenda becomes active in the pluralist model.⁹ To attain institutional status, an issue must possess the following characteristics: specificity, social significance, temporal relevance, simplicity, and categorical precedence. Cobb's argument can be best understood with the help of

FIGURE 4.2 The Pluralist Model of Agenda Setting





A CASE IN POINT

What's in a Name? The Case of the Mexican Swine Flu Pandemic

Shakespeare's Juliet famously asked, "What's in a name? That which we call a rose by any other name would smell as sweet." True, but in the world of thirty-second sound bites and coded political communication, names can have devastating consequences.

In the late spring of 2009, the news of a pandemic popularly referred to as "swine flu" swept the world. As the disease spread from country to country, the World Health Organization elevated the risk to level 4 on a five-point scale for a global pandemic, and panic set in. Although the virus turned out to be much less virulent than the seasonal flu virus that afflicts the world every year, it created problems on several fronts.¹⁰

First, the name "swine flu" created problems for the \$15 billion U.S. pork industry.¹¹ Although the genetic makeup of the offending virus contains parts of human as well as avian flu, fearful consumers began avoiding pork products despite assurance from the U.S. Centers for Disease Control and Prevention that any properly cooked pork product was perfectly safe. In the wholesale market, pork prices instantly dropped from \$70 to \$59 per 100 pounds of hog. Some small operations in agricultural states like Iowa faced economic ruin. China, Russia, and the Ukraine—some of the largest consumers of pork products—canceled imports from Mexico and parts of the United States. In Egypt, Muslim crowds attacked pig farms operated by Coptic Christians. The Christians claimed that the fear of the pandemic gave the perfect cover for the attack. In Israel, an ultra-Orthodox deputy health minister declared that his country would call it "Mexican flu" in order to prevent Jews from having to say the word "swine."

the story of Elián González, a castaway boy from Cuba. In November 1999, the six-year-old, his mother, her boyfriend, and eleven others fled the Communist-controlled island for a better life in the United States. In the middle of the ocean, their boat sank, and little Elián was the only survivor. After a fishing boat plucked him from the ocean, he was turned over to the U.S. Coast Guard, and a custody battle began to brew between Elián's father in Cuba and the fiercely anti-Castro exile community in Miami. Elián's father was his legal guardian and had full rights to take him back to Cuba, but the battle instantly became a political issue, a rallying point against the Communist regime. After the 11th Circuit Court of Appeals ordered the boy returned to his legal guardian and the U.S. Supreme Court refused hear his case, the Clinton administration (in particular Al Gore, who was preparing to run as the Democratic candidate for the presidency) was caught in the middle. The executive branch was obliged to follow the law, but politics dictated that they support the Cuban community in Florida. This political dimension was complicated by the fact that the Cuban American community has traditionally supported the Republican Party. The complex battle spurred a media spectacle that gripped the nation, but Attorney General Janet Reno ended it when she ordered that federal agents whisk the terrified child away in the middle of the night and return him to Cuba.

Mexico suffered the most serious damages from being associated with the flu. Although names are given to the area where the virus is first detected—such as the Spanish flu of 1918 or the Hong Kong flu of 1967—that does not mean that the virus originated from that exact location. Most medical historians believe that nearly all flu viruses originate in southeastern China, where animals and people live in close proximity in densely populated areas. Nevertheless, the term “Mexican flu” caused China to quarantine Mexican nationals for a prolonged period, spurring diplomatic tension between the two nations. Mexico finally sent a chartered plane to China to pick up some of its stranded citizens.

Fear can bring out the worst in us. In the United States, the “Mexican flu” gave a boost to those who were out to make political capital by spreading xenophobia against illegal immigrants from our neighbor to the south. A number of politicians and political commentators began conflating the two unrelated issues of illegal immigration and the pandemic, elevating xenophobia to a new height.

Finally, the term pandemic conjures up images of the black death of medieval times or the flu epidemic of 1918 that killed millions of people and depopulated parts of the world. Yet the definition of pandemic is simply the spread of an infectious disease within a significant proportion of the population in several countries. While the swine flu killed fewer than 100 people worldwide in a matter of a couple of months, we should keep in mind that every single day nearly 100 people die from influenza, most of whom are elderly or have immune deficiencies. The combination of the three terms Mexican, swine, and pandemic created havoc by shaping the definition of the problem and how the public viewed it. In the end, many agencies started referring to the virus by its neutral technical name, H1N1.

Discussion Points

1. Why are names important in the process of public policy making?
2. What do the problems caused by naming the flu virus say about human nature?

Although Elián’s eight-month turn in the media spotlight is now a minor footnote in history, in retrospect it might have changed the history of the world. Al Gore was defeated in Florida in 2000 by a handful of votes. If the Gonzalez saga had not pitted the Cuban American community against the Democratic administration, Al Gore might have secured a few more votes in Florida and George W. Bush might not have won the electoral vote. Let us now examine the pluralist model in light of the Elián González saga.

Specificity. This characteristic describes how broadly an issue has been defined. When advocates of an issue define it in a way that attracts wide public support, it is more likely to be picked up by the political elite. When Elián González requested asylum from the United States, his Republican supporters portrayed his cause as a fight between freedom and communism. From a different angle, it could have been seen as a simple case of a custody battle over an unfortunate little boy. However, the issue was framed in political terms, which helped inflame passion and translated into a political agenda. There are numerous examples of systemic agendas being translated into political action due to their framing (see “A Case in Point”). For instance, pro-tobacco forces depicted the debate over underage tobacco use as a battle for an individual’s right to choose, while

those in opposition framed the issue in terms of public health and, perhaps more important for some voters, in terms of money spent on combating the consequent illnesses. In both cases, by using terms that raise emotion, activists on both sides improved the chances that their issues would reach the institutional agenda. By framing them in a wide context, they attracted far greater public support than they would have if they had construed the issues narrowly—as one child's request for asylum or as a single industry's campaign to advertise its products as it wishes.

Social Significance. When advocates frame a complex issue in a way that makes sense to constituents, the issue attains a level of social significance. This simplification makes it easier for a political leader to champion the cause and for others to lend support. Similarly, when setting an agenda for the courts, especially the Supreme Court, advocates give their issue social clout by defining its broad constitutional implications. The question of whether grandparents should retain visitation rights when their sons or daughters get divorced might be resolved by a family court, unless advocates couch their argument in terms of the parents' right to make decisions for their children. Such an argument could elevate the issue to the high court. In the case of Elián González, the social significance of opposing the repressive regime was readily understood by the Cuban American community and, in fact, the rest of the nation.

Temporal Relevance. This term refers to whether an agenda has only short-term relevance or deep and enduring implications. Supporters of Elián González claimed that they fought not only for Elián but also for Cuba's future generations, to liberate them from dictator Fidel Castro. Proponents of a ban on teen smoking asserted their intention to save today's children from future addiction, illness, and death.

Simplicity. An issue improves its chance of being embraced by the political elite if it can be easily understood. Supporters of Elián González asked this simple question: should Elián be reunited with his father, or should he settle in America to fulfill his mother's last wish? When advocates frame issues in a straightforward manner, they can communicate them easily to the general public. In contrast, when they convey complex scientific information—about global warming, for example—their arguments are likely to languish in the arena of public debate unless they can be summarized in terms of costs and benefits.

Categorical Precedence. Finally, lawmakers are more likely to champion an issue that is a matter of routine legislative action—or has categorical precedence—than one that is unique and far from routine. Issues such as environmental protection and global warming tend to face greater resistance in Congress than, say, a specific military procurement measure. In the case of Elián, the fight against the Castro regime was part and parcel of the politics of the Cuban American community. Therefore, it was easy for politicians to pick up the issue and to force it through the political processes.

In his later work Cobb extended the pluralist model of agenda setting. In his original model, he had noted that agendas reach the institutional level when driven by public demand, as voiced through community leaders or the media. In his revised model Cobb observed that agendas also flow to institutions when initiated by political leaders, who rally widespread public support. Moreover, some agenda setting goes on out of public view, in the corridors of the legislative and executive branches.¹²

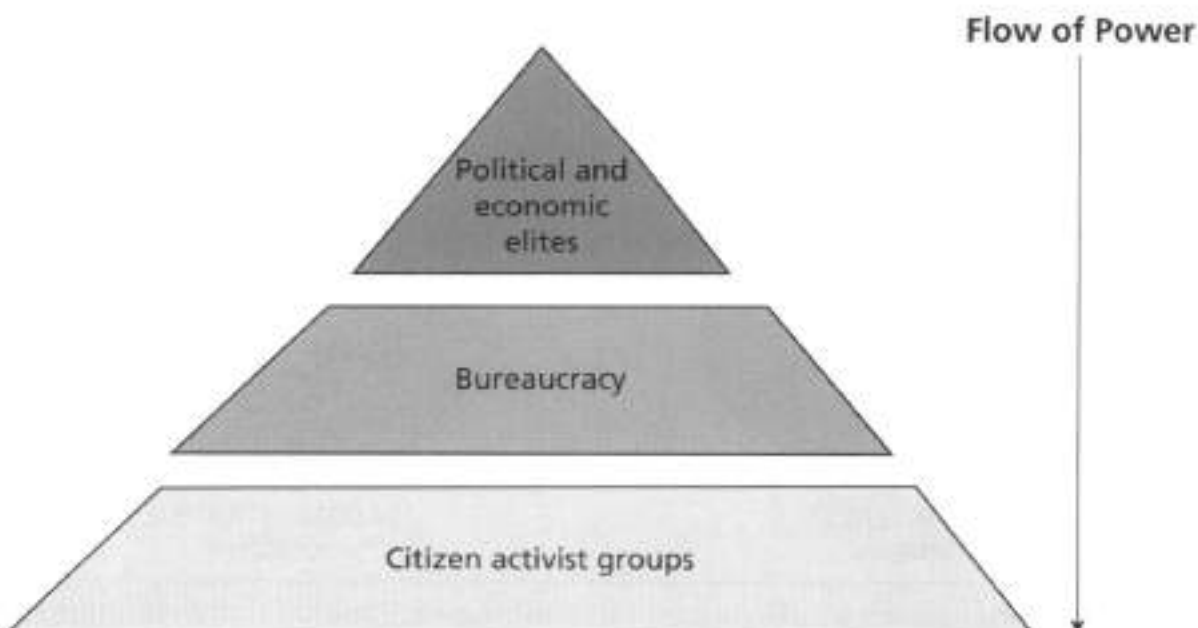
The Elitist Model

While some scholars ask *how* policy concerns reach the agenda, others ask *who* sets the agenda. Proponents of the hierarchical, or elitist, theory of agenda setting focus on the structure of decision making. They note that power is concentrated in the hands of a few. Political elites wield authority over a multitude lacking ready access to the corridors of power. The power structure resembles a pyramid, with power flowing from top to bottom (see Figure 4.3).¹³

Scholars favoring the elitist, pyramidal model adopt one of two views of agenda setting. The first view, described by decision scientists James March and Johan Olsen, points out that elites focus on issues in three ways.¹⁴ First, they select issues randomly from a wide universe of possibilities. If you are a freshman representative in the House, you can focus on an infinite number of issues. You can pick up issues here or there—perhaps military appropriations now, the environment next, and possibly welfare down the road. Second, elites choose issues in which they specialize. If you are on the Defense Appropriations Subcommittee, you tend to pay attention to the issues in your area of expertise. You focus on defense spending while ignoring or setting aside other issues. Third, when selecting issues, elites observe hierarchies. A hierarchy governs the work of the Appropriations Committee. When the committee considers an issue, your staff reviews it first. As the member of Congress, you deal only with those issues that your staff has already vetted. If you are a junior member, your area of attention differs from that of the chair of the subcommittee. Within an organization, participants focus their attention in a hierarchical way. Therefore, the process of agenda setting is inevitably hierarchical.

The second view of elitist agenda setting, advanced by political scientists Thomas Dye and Harmon Ziegler, argues somewhat cynically that society's elites select issues that serve their own interests.¹⁵ Separated from the masses, who are by and large apathetic, elites such as lawmakers, corporate representatives, and other interest groups

FIGURE 4.3 The Elitist Model of Agenda Setting

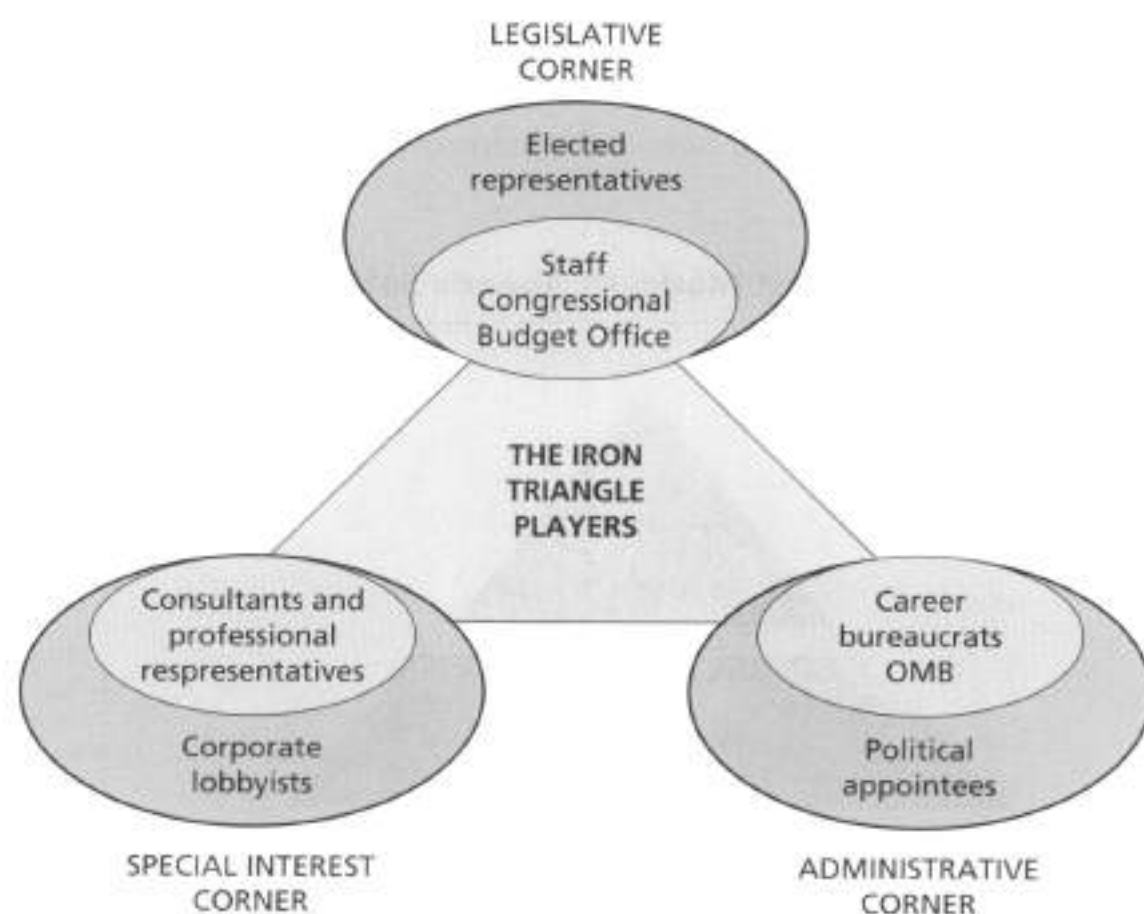


exercise overarching influence on the institutional agenda.¹⁶ Only when issues become important to elected officials or bureaucrats are they placed on the government agenda. Political and bureaucratic elites give priority not to the most salient issues but to those that best serve their own interests.¹⁷ For example, critics of President George W. Bush's energy and environmental policies argued that they protected the corporate interests, especially those of the oil, natural gas, and manufacturing industries.

Another model of agenda setting, one closely linked to the elitist model, is the **subgovernmental model** (see Figure 4.4).¹⁸ The model describes an **iron triangle**, with political elites (Congress) at the apex and, at the base, those who dominate agenda setting—administrative elites (the bureaucracy) and special interests, along with their staffs and other experts. This rigid model assumes a stable coalition of players in the policy-making process. Some scholars have proposed a model slightly different from this one, in which the coalitions are flexible and open to outsiders. Outside players such as think tanks, academics, and journalists often make significant contributions to policy processes.¹⁹ By publishing an insightful book or article, or by broadcasting a penetrating investigative report, these players can join a loose, informal coalition with policy advocates and community activists.²⁰

John Kingdon presents another model of agenda setting, one that combines features of the pluralist, elitist, and subgovernmental models.²¹ According to the **garbage can**

FIGURE 4.4 The Subgovernmental Model of Agenda Setting



From J.P. Lester and J. Stewart, Jr. *Public Policy: An Evolutionary Approach*, 2d ed. © 2000 Wadsworth, a part of Cengage Learning, Inc. Reproduced by permission. www.cengage.com/permissions.

Note: OMB is the Office of Management and Budget.

TABLE 4.1

Comparison of Pluralist and Elitist Models of Agenda Setting

	Pluralist model	Elitist model
Power	Based on size of group and its access to resources	Concentrated in the hands of a few
Centers of power	Multiple	Few
Values	Shared by masses and elites	Basic consensus among elites; values of elites differ from those of masses
Social mobility	High; elites permit input from masses and confer when making a decision	Low; masses exert minimal influence over elites
Influence	Individuals can influence elites	Individuals cannot sway elites; elites are highly insulated from apathetic masses
Outcome	Depends on many compromises among competing groups	Depends on elites' directing policy from top to bottom, serving their own interests

model, when making decisions we act like a magician reaching inside a black hat and pulling out a rabbit. That is, we make choices in seemingly chaotic fashion, in a policy-making environment of “organized anarchy.” We do not seek the absolute best decision. If we were to do so, we might spend our entire lives looking for it. Even if we did find the absolute best choice, the cost of the search might outweigh the benefits. Instead of searching, we choose the option that **satisfices**—that is, the one that is satisfactory and sufficiently good.²² According to Kingdon, we make decisions using “imperfect rationality” (satisficing behavior), incomplete preferences (preferences that cannot be placed in strict order), and fluid participation (participation without structure or hierarchy). We make decisions in a mixed environment—one characterized both by the structure of the cozy iron triangle and by the randomness of the pluralistic model, in which we participate in give-and-take based on the mobilization of groups. Recent studies present evidence of this garbage can, or semichaotic, model of agenda setting but give greater weight than does Kingdon to the model’s elitist, top-down features.

The distinguishing characteristics of the pluralist and elitist models, which present contrasting views of American politics, are listed in Table 4.1.

Impact of Sensational Events

Our discussion of how an agenda attains institutional status would not be complete if we did not mention the effects of sensational events on policymaking. In his exhaustive work, Thomas Birkland analyzed the impact of hurricanes, earthquakes, and oil spills on agenda setting.²³ He found that we are easily moved by pictures of human suffering; incensed, we sometimes demand immediate action. The devastation wreaked by Hurricane Katrina in New Orleans in 2005 provides a perfect example to support Birkland’s thesis: reports and images of the horrific conditions endured by residents of the predominantly poor Lower Ninth Ward served to thrust several issues, such as poverty and emergency preparedness, to the forefront of the national agenda. However, the most clear-cut case of a sensational event causing rapid political action is, of course, the 9/11 attacks. The sights



A CASE IN POINT

The Power of a Focus Event: Making Schools Safe

The worst nightmare of every parent came true in April 1999, when TV news programs showed terrified students streaming out of the Columbine High School campus in Littleton, Colorado. Responding to the bloodshed caused by two troubled teenagers, Eric Harris and Dylan Klebold, the federal Bureau of Alcohol, Tobacco, and Firearms (ATF), along with school boards all over the country, moved aggressively to combat school violence and prevent another tragedy. Although policymakers had debated school safety for years, the incident at Columbine brought the issue to the institutional level, and the American public demanded action.

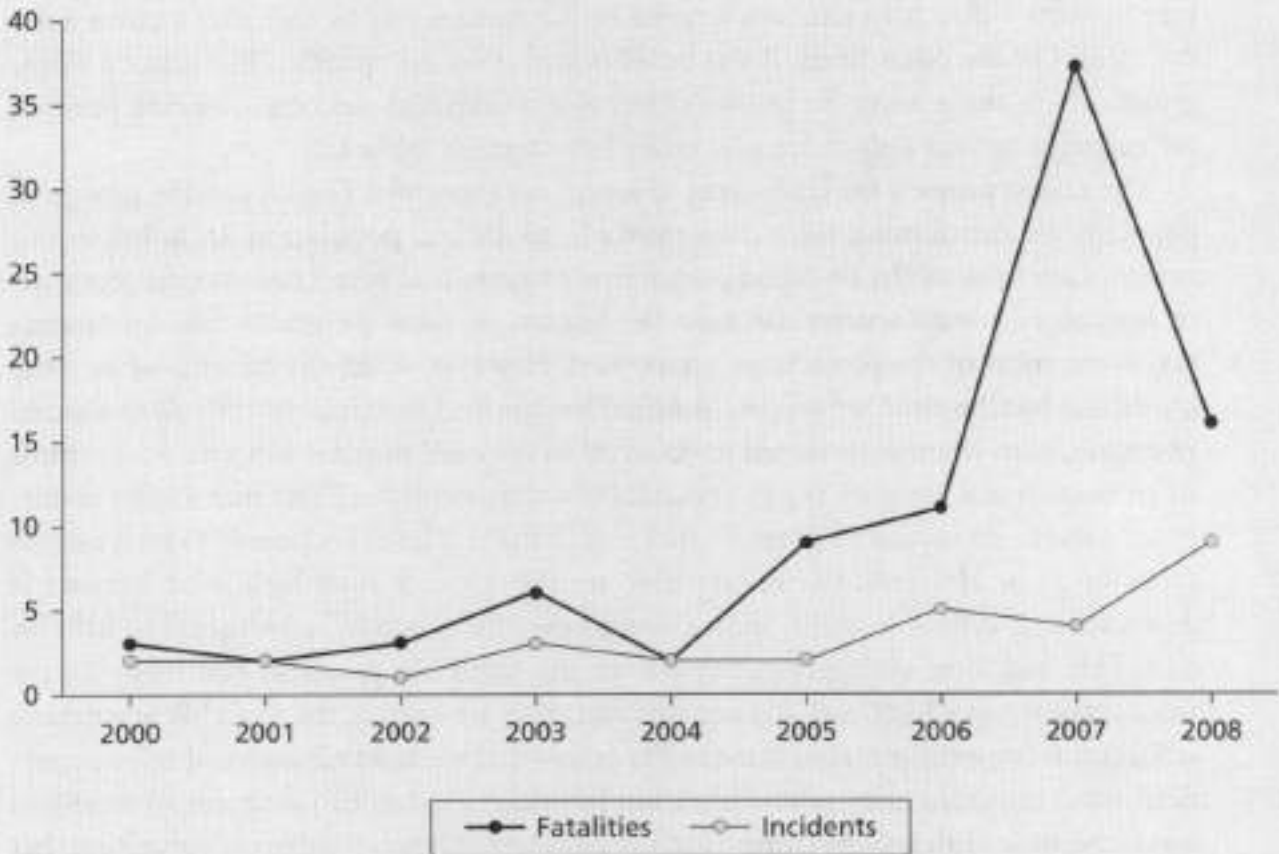
While many schools responded by installing expensive metal detectors at school entrances, the ATF worked with a threat assessment company to help school administrators identify troubled students who might be on the verge of committing violent acts. The controversial program, known as Mosaic 2000, tested students in twenty schools—primarily high schools—as part of a national pilot study.²⁴ Mosaic identified potentially violent students at Reynoldsburg High School, in a suburb of Columbus, Ohio, by asking students 150 questions crafted by experts in law enforcement and psychiatry. Administrators also asked parents, friends, and school counselors to evaluate students. The students were then rated on a scale of one to ten.

Needless to say, the Mosaic study provoked controversy. At Reynoldsburg High, school administrators, parents, and the school board favored the program, as did Ohio's attorney general. However, the Ohio chapter of the American Civil Liberties Union expressed concern that the program breached students' civil rights, including their right of confidentiality. The group called Mosaic a "technological Band-Aid" that fed on parents' fear. Critics also questioned the program's effectiveness in identifying potentially violent students. At Columbine, the perpetrators were good, though troubled, students with access to guns, students who unexpectedly snapped because they felt victimized by bullies and the school system. In a recent book, journalist Dave Cullen pointed out that although Harris probably was a sadistic psychopath and Klebold was clinically depressed, there was no specific characteristic that would have identified the two young men before the incident took place.²⁵ When such students are involved, is Mosaic just as helpless as school administrators in identifying the potential assailants?

Similar to many other sensational events, the Columbine massacre created its own copycats. As shown in the graph below, incidents of school (and college) shootings took more lives in subsequent years. In addition to a rising trend in the United States, other nations, such as Germany and Finland, have also seen senseless killings by young people.

of the collapsing twin towers and hapless people jumping to their deaths have been seared in our collective memory ever since. Terrorist organizations know the importance of sensational events in shaping a nation's agenda; as a result, Alex Schmid and Janny de Graaf have argued that these acts are forms of political communication.²⁶ For the past forty years, international terrorist acts have killed about 370 people per year—while this is a deplorable amount, it is still slightly fewer than those who drown in bathtubs in the United States every year.²⁷ Because terrorist attacks generate such a spectacle, though, they force governments to overreact. Thus, President Bush felt the need to respond to the

School Shootings in the United States



Discussion Points

1. How did school safety reach the institutional agenda?
2. How did the Reynoldsburg school board respond to public concern? Do you think their response was effective?

9/11 attacks in a dramatic manner, quickly declaring a global war on terror and launching a two-front war in Afghanistan and Iraq.²⁸

However, sometimes it takes more than shocking pictures of disaster to force an issue from the systemic to the institutional agenda. Birkland argues that an agenda can be activated by a **focus event** that spurs a community to organize and demand legislation (see "A Case in Point"). The most effective community is a coherent group of professionals, experts, and activists—policy entrepreneurs—who mobilize and initiate a plan of action.

Evaluating Costs and Benefits

Although the process by which policymakers give a systemic agenda official status is complex, we can develop a simple model that explains which agendas are likely to make the official list. We begin by evaluating systemic agendas in terms of their relative costs and benefits. If the benefits and costs are diffuse, they fall on a large number of people who do not belong to a well-defined group. Most environmental projects provide diffuse benefits—they help almost everyone in the community by fostering a clean environment. On the other hand, if the benefits and costs are specific, they affect a single group, which alone bears the project’s favorable or harmful outcome. Hunting licenses, for example, benefit only those who enjoy hunting (see Table 4.2).

The easiest projects for lawmakers to accept are those that benefit specific groups of people while distributing their costs over a large, diffuse population. Included in this category are most of the so-called pork barrel projects that help a few specific segments of legislators’ constituencies. Because the burden of these programs falls on faceless taxpayers, most of the projects go unopposed. However, when the benefits of an issue are diffuse but the costs are specific, political leaders find it extremely difficult to allocate resources, even when encouraged to do so by widespread popular support. An example of an issue in this category is gun control. Polls consistently indicate that a solid majority of Americans support a measure that would make it hard for people to own certain types of guns. However, lawmakers have trouble passing such legislation because it promises only diffuse benefits, lacks a single powerful group of constituents to fight on its behalf, and faces strong opposition from the small but powerful gun lobby. In the third category, in which both the benefits and costs are diffuse, the most likely outcome is inaction. For example, almost everyone agrees that we need educational reform, particularly at the elementary school level. But lawmakers struggle to pinpoint winners and losers, because both groups come from the citizenry at large. It is hardly surprising that for all their rhetoric, legislators pass few measures in this category. The final category comprises measures with clear winners and losers. These bills address the most controversial redistribution issues. Examples of such measures include bills that support or oppose a new airport, pit environmentalists against the timber industry in the Pacific Northwest, or redress past injustice through programs such as affirmative action.

Having explored agenda setting in terms of benefits and costs, we are prepared to briefly discuss the two ways people commonly assess the benefits and costs of a policy. Those who favor a narrow view hold that when considering costs and benefits, people strive only to maximize their own short-term self-interest. If the potential benefits for a

TABLE 4.2 Outcomes for Diffuse and Specific Costs and Benefits

Costs (losers)	Benefits (winners)	
	<i>Diffuse</i>	<i>Specific</i>
<i>Diffuse</i>	Inaction	Likely acceptance
<i>Specific</i>	Likely rejection	Conflict

group are diffuse, the group is likely to lose, because no one in the group will be willing to bear the cost of a policy that will benefit the entire group, whether or not the other members share the costs. Mancur Olson called this situation the free-rider problem, asserting that “rational” people do not want to participate in a collective action when they can get something for nothing (see chapter 1). Olson pointed out that as group size increases, the problem becomes more pronounced.²⁹ In contrast, those who advocate the broad view contend that people are motivated not just by short-term self-interest but also by commitment,³⁰ group utility,³¹ and ideology.³² Supporters of this theory assert that if people defined benefits and costs as narrowly as the self-interest camp says they do, they would never embrace the most challenging causes—from improving the environment to supporting national liberation movements. Because environmental policies benefit so many of us, we find it almost impossible to answer the question, why should I give my time to support these issues? Indeed, we cannot answer this important question if we assume that people follow only their narrowly defined self-interest. However, if we accept the premise that people often look beyond their own selfish needs and are inspired to contribute benefits to a larger community, we can explain a wide variety of altruistic behavior. When group members define an issue’s benefits and costs as “specific” to them, they may fight to keep it in front of the political elite and get it onto the institutional agenda. By taking this approach, volunteers for Mothers Against Drunk Driving spurred the passage of many local laws that got tough on drunk driving.

POLICY FORMULATION

Once an agenda attains institutional status, specific policies must be formulated. Legislators outline the course of action that the policy will follow. This plan can then be enacted into law. To determine the plan, lawmakers specify objectives, identify policy alternatives, and adopt the one alternative that gives the best results. Although analysts may participate in each stage of the policy cycle, their involvement is most critical in the phases of policy formulation, implementation, and evaluation.

Suppose we want to achieve a completely drug-free society—a desirable objective, though highly improbable. If we set an impossible goal, our policies, no matter how effective, will fail. Therefore, we start with the goal of reducing illegal drug use by, say, 10 percent in the next five years. How do we achieve this policy goal? We attack the problem by trying to curtail either the demand for drugs or their supply. We may choose to adopt one or more of the policy solutions outlined below. These policies may be enacted through legislation, executive orders, ordinances, or judicial decisions.³³

Policies tend to provide either positive or negative incentives. Many drug users want to quit but cannot do so because of the high cost of treatment. To provide a positive incentive for quitting, the government may open rehabilitation centers for treating users free of charge. As a further inducement, the government may protect the privacy of those receiving treatment. On the flip side, negative inducements include stiffer penalties for drug possession and mandatory sentencing guidelines for repeat offenders. The government also may empower law enforcement agents to confiscate property gained through drug trafficking.

Another option available to us is to modify behavior by changing the rules. We might pass laws to make detection easier or to authorize locker searches. Students’ fear

of getting caught would raise the cost of using illicit drugs and might discourage them from bringing drugs to school. Moreover, we might decide to alter behavior with educational campaigns. We could distribute factual information on the physiological and psychological impacts of drug abuse. Such information might take the form of advertisements, leaflets, or curricula on drug abuse, such as the Drug Abuse Resistance Education (D.A.R.E.) program for children.³⁴ We could also pass laws to give people rights that protect them from injuries or injustice related to drug enforcement and sentencing. Finally, we might accomplish the aims of our institutional agenda by shifting power over the agenda to a specific agency. In one of the most controversial anti-drug proposals, the federal government considered engaging the air force to intercept low-flying airplanes crossing the U.S.-Mexican border. It also authorized the navy and the Coast Guard to seize high-speed boats and ships dropping off drugs.

Explaining Behavior

Before we formulate policy and do everything necessary to achieve our policy goals—offer incentives, change the rules, distribute factual information, give people rights, or transfer power—we first must determine if our proposed solution will be effective. To make this determination, we develop a behavioral model, making certain assumptions about human nature.

If we educate potential users about the ill effects of drug use on their bodies and minds, will we significantly affect their behavior? If so, why? The answer to this question will help us develop a fuller understanding of why people become addicted in the first place, and this information will shape the policies we propose. We might want to know where our target population is most likely to get its information. That way, if we conduct an advertising campaign, we will know from our behavioral model what the ads should say and what kind of ads to run.

To check our behavioral model we can use statistical techniques, testing our hypotheses with existing, **secondary data** or with data collected by the researcher, known as **primary data** (see the discussion of statistical techniques in chapters 7 and 9; see also information on data gathering in chapter 8).

Forecasting Effects

We can use our behavioral model to make a forecast. Scientific forecasting is not fortune-telling. The primary distinction between the two is that although fortune-telling makes a prediction about a single unique event, answering a question with a definitive yes or no, scientific forecasting takes place in the context of probability. When a scientist forecasts a 99 percent probability and the event does not take place, the forecast may still be valid. Embedded in the forecast is a small but definite 1 percent chance of the event's not taking place. What the scientist intends to say with the forecast is that if the same conditions prevail over many tries, the event will occur 99 percent of the time. In forecasting we can never develop a model that fully eliminates uncertainty.³⁵

Scientists base their forecasts on explicit or implicit assumptions about people's behavior, assumptions that they test by using control variables. Similarly, policymakers manipulate control variables to achieve their stated goals. If we assume that a rational human being is more inclined to abstain from drugs if given enough information about

their harmful effects, such information becomes our control variable, or the factor designed to induce certain behavior. Forecasting is often like looking in the rearview mirror while driving a car. If we want to understand the future, we must study the past. We make most forecasts using past data. If the road behind is the same as the one ahead, we are in a good position to forecast our future. However, when the road changes, curving and zigzagging, our forecasts lose their accuracy (see the discussion of forecasting techniques in chapters 10–12).

After setting goals and identifying alternatives, we pick the best possible option. We make our choice using cost-benefit or cost-effectiveness analysis (see chapter 14).

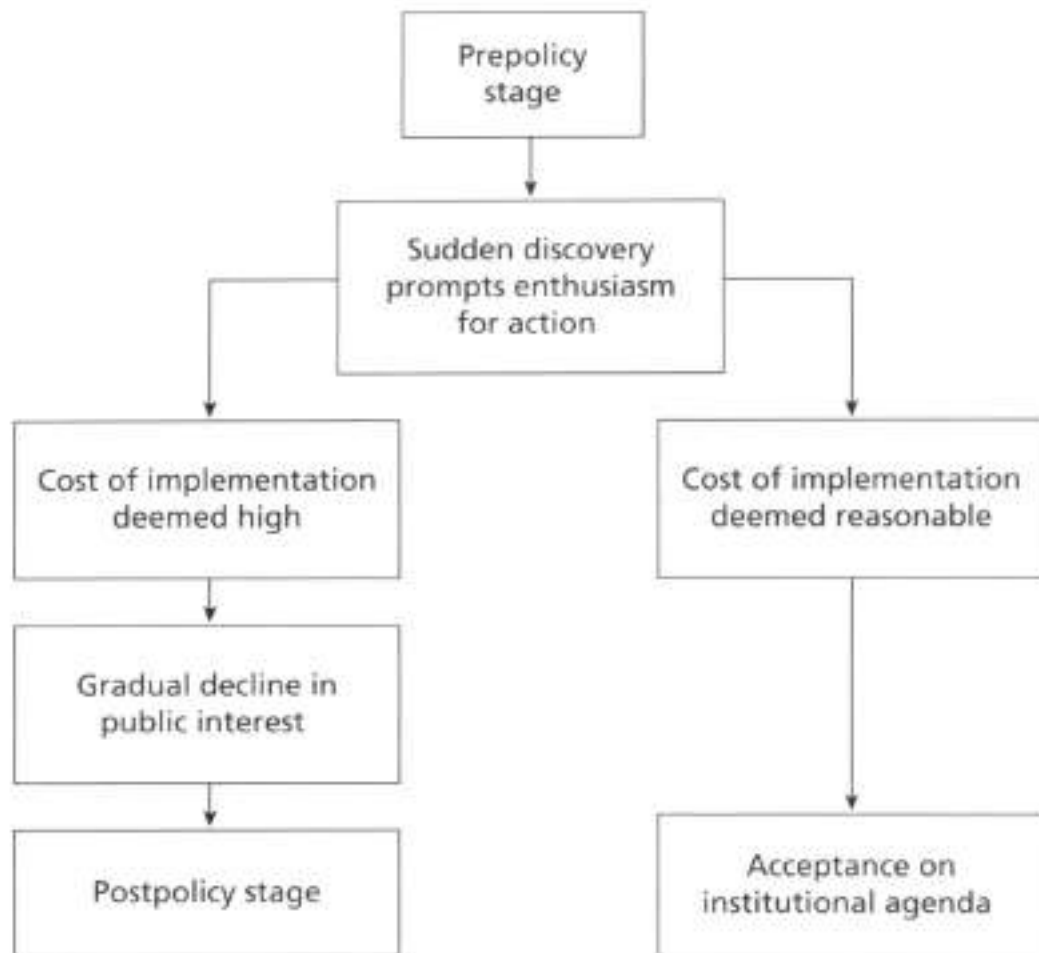
Conflict, Inaction, and Nondecision

We define public policy as what the government decides to do and not to do. The government can make as loud a statement by refusing to get involved as by taking action. Many agendas, from poverty to nuclear weapons, swirl around in society and are debated on college campuses, discussed in community halls, and written about on editorial pages. However, the vast majority of them do not make it to the institutional agenda. The government and the power elite may actively prevent some of the agendas from ever reaching the institutional stage, either by force or through deliberate neglect. Government officials in the South did not make a concerted effort to address black poverty and discrimination until civil rights legislation compelled them to take action. Authorities frequently used brute force to block attempts to discuss the relevant issues.³⁶

The government, however, does not need brute force to keep certain agendas from reaching institutional status. Most often, it blocks consideration of agendas unpopular with the elite simply through benign neglect. Why does it prevent some agendas from achieving institutional status? Evidence suggests two reasons for such failures (see Table 4.2). First, some issues provoke huge conflicts in society, especially when these issues identify specific winners and losers. If passions run high on both sides, political leaders often try to sidestep the issues. The primary goal of politics is managing conflict. No organized society can survive without suppressing certain disagreements that can develop into huge conflicts. Take, for instance, the issue of gun control. Suppose the governor of your state fervently believes that banning weapons is the only way of reducing violence. If the governor's party controls the legislature, the governor may feel free to take a radical stance on the issue. However, if others in the state passionately defend their right to own guns, the clash between the two viewpoints may lead to open conflict. Facing this ominous prospect, the political elites of both parties may refrain from pushing the agenda forward.³⁷

In addition, if both the winners and the losers are diffuse, policymakers have no incentive to elevate an agenda from the nebulous systemic zone to the institutional stage. Many issues with merit never achieve institutional status. Thus, despite their scientific relevance, numerous environmental issues of global proportion remain unaddressed.

On the other hand, sometimes an issue that has been neglected for years suddenly commands the attention of the general public and the power elite. Adapting the arguments of social scientist Anthony Downs, we can show how an issue goes through an

FIGURE 4.5**Decision and Nondecision in Policy Formulation**

“attention cycle.” As it proceeds through the cycle, it either fades from the public agenda or pushes toward institutional acceptance (see Figure 4.5).³⁸ At the systemic, prepolicy level, an agenda remains dormant. If a sudden discovery—such as news about the spread of the AIDS virus, the existence of an ozone hole over the Antarctic, or the alleged security leaks at Los Alamos National Laboratory—mobilizes influential members of the policy community, they may issue an enthusiastic call for action. However, if they perceive an unacceptably high cost, the agenda may lose public interest and drop into “a twilight calm of lesser attention and spasmodic recurrences of interest.” Downs suggests that such agendas have three things in common: they affect numerical minorities (for example, poor people, homeless people), they advocate social arrangements providing benefits (such as high-speed rail transport or new subway systems) to a majority or a powerful minority, and they no longer captivate the public with exciting ideas (for example, the space program’s lunar landing missions—routine flights to the moon did not sustain initial public fascination).

If influential policymakers find the costs to be reasonable, a sudden discovery may motivate them to transform a long-standing agenda into an accepted public policy. In the late 1990s, advances in DNA analysis brought capital punishment to the forefront of public awareness. A shocked public learned that new methods allowing researchers to

read an individual's genetic code had proven the innocence of a number of inmates slated to face capital punishment. In 1999 Illinois governor George Ryan, a Republican, suspended executions until he could be sure that only the guilty would be put to death.³⁹ In the Senate, Patrick Leahy, D-Vt., and Gordon Smith, R-Ore., jointly proposed that defendants be guaranteed competent lawyers and provided with DNA testing that might prove their innocence. Some states, including California, made similar proposals. Several religious leaders and anti-capital punishment advocates called for a moratorium on capital punishment. Once nearly forgotten, the movement questioning the appropriateness of capital punishment quickly returned to the public consciousness. Foes of capital punishment drew strength from a steadily declining crime rate during the past decade, which enhanced the public's willingness to accept shorter sentences. Because policymakers saw as reasonable the cost of making DNA testing mandatory and abandoning capital punishment, the policies quickly gained the government's attention.

POLICY ADOPTION

Policy adoption comes at the end of the formulation process. Once policy analyses are complete, they are sent to the official decision-making bodies, which enact legislation or issue executive orders. In the long cycle of policy formation, the adoption stage is for the most part well-defined—if not at the beginning, then certainly at the end of the stage, when the legislature takes a vote or the chief executive issues a formal executive order. Typically the analyst's job ends with policy formulation, when decision makers review the alternatives and analyses of their respective impacts.

In a policy analyst's ideal world, his or her recommendations will be seriously discussed, and policymakers will pick from the alternatives he or she has identified. However, in the real world, lawmakers and the chief executive may be guided by a welter of conflicting concerns, of which the analyst's recommendation is but one. As the formal decision stage approaches, the outcome may reflect a long history of bargaining and deal making based on the decision makers' values, their constituents' concerns, their party affiliations, public opinion, and pressure from special interest groups.⁴⁰

When we look at how decision-making bodies work, we may become cynical about the motives of elected officials. In opinion polls the American public repeatedly expresses its skepticism about Congress's ability to lead the country.⁴¹ The question is, when deciding policy issues, do elected representatives vote only their self-interest (defined as the interest of their constituents and their political action committees), or are they also guided by "ideology"? In an interesting study, researchers Joseph Kalt and Mark Zupan followed legislation through the adoption stage of the policy cycle.⁴² They evaluated the voting records of U.S. legislators debating the Surface Mining Control and Reclamation Act (SMCRA), which resulted from a protracted political struggle. Congress passed two versions of the act, in 1974 and 1975. President Gerald Ford vetoed them on both occasions. President Jimmy Carter finally signed the SMCRA into law in 1977. The act required that strip-mined land be restored to its pre-mining state. It also established the Abandoned Mine Reclamation Fund and clarified previously undefined rights to water and land in strip-mined areas.

The self-interest theory says that lawmakers vote only the economic interests of their constituents, with a view toward winning reelection. However, the survey

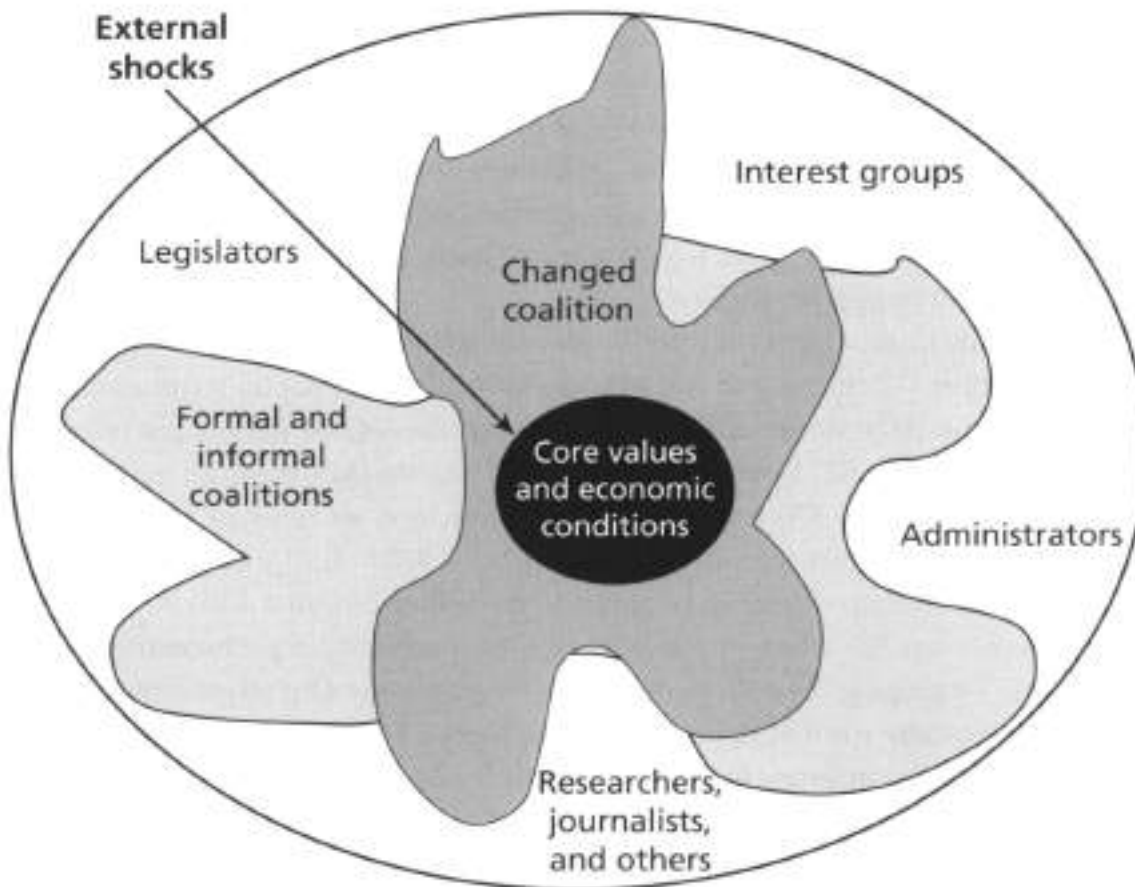
indicated that self-interest, as measured by the importance of mining in a lawmaker's state or district, does a poor job on its own of explaining voting records. The theory was a better predictor of voting behavior when researchers considered not only self-interest but also ideology, as measured by the League of Conservation Voters, an ideological watchdog. The survey indicated that when voting on legislation, U.S. senators and representatives are guided by both self-interest and ideology. Several other studies found similar results.⁴³

POLICY IMPLEMENTATION

After a public policy has been adopted, it must be implemented. Adopted policies, particularly legislative acts, almost never specify exactly what is to be done. If a city council passes an ordinance prohibiting skateboarding on public sidewalks, it rarely tells administrators how to implement the measure. Administrators have numerous options. To understand the complexity of their choices, consider the following facts. Because it is impossible to patrol every street corner at all times, the city police department must decide where and when to enforce the new law. Police must decide whether to monitor only those streets with a high volume of pedestrian traffic or to patrol both these areas and those where young people are active, such as parks, schools, and beaches. They also must decide whether to patrol on foot, on bicycles, or in cruisers, which are fast but inefficient in crowded streets. Even after making these decisions, administrators must figure out what fraction of the police department's resources should be devoted to enforcing this new law. In some areas, the new mandate might come with a specific allocation of money, but that is not always the case. Because the police department might prompt legal challenges when enforcing the new law, officers must try to strike a balance between the rights of individuals and the demands of cities.

During the past fifty years, scholars have devoted a great deal of energy to understanding the major players and coalitions in agenda setting and policy formulation. But what happens after the laws have been passed? Do the formal and informal coalitions developed in the previous stages of the policy cycle disappear in the implementation stage? Or does the power coalition that brought about policy change remain active in overseeing implementation? If the power coalition continues to shape implementation, does the structure of the original coalition remain the same, or does it change?

In a series of studies conducted primarily in the 1990s, policy analysts made a number of useful observations about the implementation stage.⁴⁴ Figure 4.6 illustrates their findings. As shown in the figure, institutional agendas are set and policies formulated by a loose coalition that may include several formal as well as informal players, who cluster around a set of "core values." According to Paul Sabatier and Hank Jenkins-Smith, these core values are the glue that holds together a disparate coalition.⁴⁵ This belief system is highly resilient and often resists changes, particularly by members of the advocacy groups in the broad coalition. Without any strong external shocks, these coalitions tend to remain stable throughout the implementation process. However, when economic conditions change, new information comes to light, or unexpected events occur—such as legal challenges or technical glitches—the

FIGURE 4.6 Impact of External Shocks on Policy Coalitions

composition of these coalitions may alter significantly. For example, Brian Ellison studied the Animas–La Plata River project in Colorado, near the New Mexico border.⁴⁶ Originally seen as a water project, it enjoyed the support of a pro–water development coalition. However, when a government study questioned the financial feasibility of the project and noted that it could jeopardize the river’s endangered squawfish population, these external shocks changed the configuration of the coalition of Indian tribes, environmentalists, and probusiness groups, thereby altering the project several times during implementation.

POLICY EVALUATION

If there is one area over which policy analysts exercise primary responsibility in the policy cycle, it is the evaluation phase. Analysts inquire about the possible impact of an adopted policy. Does the policy meet the greater needs of society, and is it achieving its goals? The first question is broader than the second. We hope that the goals of society and those of the policy are compatible, but they sometimes conflict with each other. Suppose your state government has initiated a new education policy to boost school performance. It has allocated funds to prepare students for national standardized achievement tests. If at the end of a certain period the average test score has improved,

we may conclude that the program has achieved its goals and that the money has been well spent. However, a broader question arises if we look at the total impact of the program, particularly on children from traditionally disadvantaged and non-English-speaking families. Suppose that while average test scores have improved, the program's curricular changes have devastated less prepared students, increasing dropout rates. In addition, suppose that legislators funded this program at the expense of another project, a state-funded special initiative for children with learning disabilities. We might find that the opportunity cost of such a program exceeds the benefits. When we evaluate the policy on its stated goals, we get a much more positive picture than when we analyze its overall impact on society.

We may evaluate the impact of a policy on several levels. For example, we may examine **policy output**.⁴⁷ Returning to our example of traffic cameras on street corners, we note that in June 2000, traffic officials in San Diego cited 4,000 drivers per month for violations, fining them \$271 per ticket.⁴⁸ Lockheed Martin IMS, the local manufacturer of the cameras, received \$70 from each ticket. Therefore, we estimate that the policy output equals the city's net revenue per ticket ($\$271 - \$70 = \$201$). At this rate, the city expects to make annual collections of about \$9.65 million ($4,000 \times \201×12).⁴⁹

We also measure the effect of public policy by conducting a **performance evaluation**—that is, by gauging the policy's impact on target groups. Our target group includes drivers who typically run red lights, causing accidents. However, performance evaluation does more than measure how many people in the target group are affected by a program; it also evaluates the program's total effectiveness. Analysts measure the performance of the cameras by looking at accident data and conducting research to see if the devices reduced the total number of accidents.

In addition, policy analysts may want to evaluate the impact of a policy from an even broader perspective than the one afforded by performance evaluation. In that case, they look at **policy outcomes**. The effects of a policy may not be confined to the target group or even to the immediate program objectives. Instead, they may spill over into other areas of social behavior and social conditions. Because of the traffic devices, motorists may drive carefully not only where cameras are located but also everywhere else they drive, making areas beyond the monitored intersections safer for both pedestrians and drivers.

Finally, analysts measure policy impacts by evaluating feedback. A number of motorists complained that the cameras were unsafe. The drivers argued that if they were to stop suddenly to avoid hefty fines, they might cause accidents. They also argued that at night the blinding flashes from the cameras caused confusion, which led to accidents.⁵⁰ This kind of feedback helps analysts change policies to better fit objectives.

POLICY CHANGE

As alluded to above, policies evolve and go through numerous changes until some of them finally are terminated. If we follow the life of any single public policy, we find that it is constantly shaped and reshaped by at least seven factors:

1. Incremental changes in the dynamics of society smooth out the rough edges of implementation.
2. New statutes contradict or invalidate parts of the existing policy.

3. Lawsuits or questions of constitutionality challenge the policy in court.
4. New technology alters the feasibility of the policy.
5. New discoveries or revelations change public support for a program.
6. Political and economic circumstances change, imposing different conditions on an existing policy.
7. Elections cause a major ideological shift that changes the policy.

In other words, as the world around us changes, so do all major policies.

If we look at capital punishment, we see how society's attitudes have changed over the years. A strict "eye for an eye" attitude of swift justice has given way to a more reasoned approach, one that considers carefully the social effects of such a policy. The history of capital punishment goes back to the earliest period of recorded penal codes. The Code of Hammurabi (1790–1750 BCE) recommended capital punishment for twenty-five different crimes.⁵¹ In the seventh century BCE, Draco, the Athenian law-giver, from whom we get the term *draconian*, prescribed death for every kind of offense. It is interesting to note that Sir Thomas More explained in *Utopia* that in his imaginary land, where pure reason guided all public policy, capital punishment for most offenses, such as stealing, was prohibited. This is because if thieves knew their punishment, they would have every incentive to kill anyone wanting to turn them in. In contrast, capital punishment was prescribed for anyone discussing politics outside the Senate.⁵² Brought to the new world by European colonists, capital punishment was liberally administered for high crimes against the state as well as for petty theft. Soon, though, a movement to abolish capital punishment gained ground in the United States. As governor of Virginia, Thomas Jefferson recommended capital punishment only for treason and murder, but after a stormy debate, the legislature defeated the bill by one vote. In 1972, the Supreme Court, reviewing three separate cases, struck down capital punishment in several states and set the standard for capital punishment. Under this guideline, the Court would consider capital punishment "cruel and unusual" if it (a) was too severe for the crime, (b) was arbitrary (that is, not uniform—applied to some but not to others in similar situations), (c) offended society's sense of justice, or (d) was less effective than another punishment.⁵³ However, during the next seventeen years, the pendulum of the Court's position on the issue swung back hard and fast. In a series of challenges, the Court upheld capital punishment in a number of state cases.⁵⁴ When opening old cases in the 1990s, several state courts started admitting DNA evidence. Capital punishment remains an important issue today, as demonstrated by the repeal of capital punishment in New Jersey and New Mexico in 2007 and 2009, respectively.⁵⁵

Policy scientists note four patterns of policy change:

1. **Linear**, when one policy is replaced by another.
2. **Consolidated**, when programs with similar goals are combined.
3. **Split**, when an agency enforcing a certain policy grows too large to administer the policy effectively. The agency is broken into smaller components, and each of them receives part of the original policy.
4. **Nonlinear**, when social conditions change or technological advances generate new information, prompting drastic or major policy changes.

POLICY TERMINATION

As the government accumulates trillions of dollars of public debt more than sixty years after John Maynard Keynes recommended to President Franklin Roosevelt that the government increase spending during recessions and reduce spending during periods of inflation, some policy scientists charge that Keynes ignored political realities. They note that government spending has been elastic in only one direction: upward. Government officials have found it easier to increase spending or propose a new public policy than to terminate one after it has been implemented. As Garry Brewer and Peter deLeon point out, the cavalry did not gallop into the sunset with the introduction of mechanized regiments; the March of Dimes did not stop marching after the invention of the polio vaccine.⁵⁶

Scholars define policy termination as “the deliberate conclusion or cessation of specific public sector functions, programs, policies, or organizations.”⁵⁷ They have identified four types of termination: functional, organizational, policy, and program.⁵⁸

- **Functional termination** is sweeping. If the government decided to abandon its educational responsibilities, that decision would fall in this category. This termination would cause the complete privatization of education, with the private sector managing the entire field.
- When the government does not want to end its involvement in a field but does wish to withdraw support for an agency, it engages in **organizational termination**. Although nobody expects the government to drop its educational duties entirely, many conservatives in U.S. politics have urged the federal government to end its involvement in the field by eliminating the Department of Education.
- **Policy termination** takes place when lawmakers abandon a specific policy to address changing social needs. In the American South legislators discarded discriminatory Jim Crow laws to respond to changes in the post-civil rights era.
- **Program termination** occurs when the government ends a specific program. For example, in 2006 the Bush administration ended the State Planning Grant Program, which provided funds for states to collect data and develop strategies for dealing with people without health insurance, because it was found to be ineffective and duplicative of other programs.

Why is policy termination a problem for government? Lawmakers adopt a specific public policy to address a specific human problem. They rarely solve a problem completely. Therefore, the policy lingers. It goes through many changes, but the essential goals of the policy remain intact. When legislators terminate a program, they must contend with the losers. Lawmakers' failure to distribute the losses effectively causes many of these programs to continue. This is particularly true when the costs of termination fall on a small, specific group (employees working at a military base, for example, and those who depend on the business generated by the base). It is equally true when the benefits remain diffuse (savings for U.S. taxpayers). A cynical observer would argue that every government program creates its own constituencies. These

entrenched interests prevent political leaders from taking the hard step of terminating a policy, program, or agency. Many military bases were established in the United States to fight Native Americans during the early days of the nation's settlement. Yet most of those bases remain in use even a century after their original purpose was fulfilled. Anytime lawmakers try to close a military base, the agencies near the base lobby Congress to keep it open.⁵⁹ This pattern repeats itself when political leaders debate whether to end other government programs. As a result, few programs are ever completely terminated.

KEY WORDS

Consolidated policy change (p. 95)	Policy cycle (p. 75)
Elitist model (p. 77)	Policy formation (p. 75)
Focus event (p. 85)	Policy outcomes (p. 94)
Functional termination (p. 96)	Policy output (p. 94)
Garbage can model (p. 82)	Policy termination (p. 96)
Institutional, or governmental, agenda (p. 76)	Primary data (p. 88)
Iron triangle (p. 82)	Program termination (p. 96)
Linear policy change (p. 95)	Satisfices (p. 83)
Nonlinear policy change (p. 95)	Secondary data (p. 88)
Organizational termination (p. 96)	Split policy change (p. 95)
Performance evaluation (p. 94)	Subgovernmental model (p. 82)
Pluralist model (p. 77)	Systemic, or noninstitutional, agenda (p. 76)

EXERCISES

1. Every newly elected president goes to the White House promising to alleviate poverty and economic inequality. Yet in the end, many research and policy debates produce only marginal changes in this area. After reading through the Government Accountability Office report on poverty in the United States in Appendix A, write an essay explaining why it is difficult to make radical changes in public policies. Specifically, why is it challenging to devise a comprehensive program to deal with poverty?
2. Despite recognition of the fact that we should try to substantially reduce poverty in America, the issue remains confined within the broader systemic agenda. How do you explain this in view of our discussion of agenda setting? Is there anything that can be done to bring the issue of poverty to the forefront of political debate?
3. What is the importance of interest group coalitions in public policy formation? Take an important policy that has recently been adopted and determine whether the coalition supporting it has been stable or whether it has changed over time.
4. Write an essay on the role of quantitative analysis—and those who use it—in public policy formation. Given the political nature of the process, should we look to economic analyses and statistical techniques for answers?

Notes

1. Jennifer Hanrahan, "Motorists Taking San Diego's Red-light Cameras to Court," *San Diego Union Tribune*, May 7, 2000.
2. James E. Anderson, *Public Policymaking: An Introduction*, 3rd ed. (Boston: Houghton Mifflin, 1990), 78.
3. See, for example, Nick Mabey, Stephen Hall, Clare Smith, and Sujata Gupta, *Argument in the Greenhouse: The International Economics of Controlling Global Warming* (New York: Routledge, 1997).
4. James P. Lester and Joseph Stewart Jr., *Public Policy: An Evolutionary Approach*, 2nd ed. (Belmont, Calif.: Wadsworth, 2000), 65–86.
5. Roger W. Cobb and Charles D. Elder, *Participation in American Politics: The Dynamics of Agenda-Building*, 2nd ed. (Baltimore: Johns Hopkins University Press, 1983).
6. *Ibid.*, 85.
7. Anderson, *Public Policymaking*, 3rd ed., 83.
8. David B. Truman, *The Governmental Process: Political Interests and Public Opinion* (New York: Knopf, 1951). See also Robert A. Dahl, *Who Governs? Democracy and Power in an American City* (New Haven, Conn.: Yale University Press, 1961).
9. Roger W. Cobb and Charles D. Elder, "The Politics of Agenda Building," *Journal of Politics* 33 (1971): 892–915; Cobb and Elder, *Participation in American Politics*; Roger W. Cobb, Jennie-Keith Ross, and Marc H. Ross, "Agenda Building as a Comparative Political Process," *American Political Science Review* 70 (1976): 126–138.
10. Keith Bradsher, "The Naming of Swine Flu, a Curious Matter," *New York Times*, April 29, 2009.
11. Mike Baker and Michael J. Crumb, "Name Given to Virus Pains Pork Industry," Associated Press, May 1, 2009.
12. Cobb, Ross, and Ross, "Agenda Building."
13. Thomas R. Dye and L. Harmon Ziegler, *The Irony of Democracy: An Uncommon Introduction to American Politics*, 5th ed. (Monterey, Calif.: Duxbury, 1981). See also Thomas R. Dye, *Understanding Public Policy*, 9th ed. (Upper Saddle River, N.J.: Prentice Hall, 1998).
14. James G. March and Johan P. Olsen, *Ambiguity and Choice in Organizations* (Bergen, Norway: Universitetsforlaget, 1976).
15. Dye and Ziegler, *The Irony of Democracy*.
16. James E. Anderson, *Public Policymaking* (New York: Praeger, 1975).
17. Barbara J. Nelson, *Making an Issue of Child Abuse: Political Agenda Setting for Social Problems* (Chicago: University of Chicago Press, 1984).
18. Douglas Carter, *Power in Washington* (New York: Random House, 1965). See also J. L. Freeman, *The Political Process: Executive Bureau-Legislative Committee Relations*, rev. ed. (New York: Random House, 1965).
19. For an excellent, detailed discussion of how conservative think tanks have shaped public policy, see Trudy Lieberman, *Slanting the Story: The Forces That Shape the News* (New York: New Press, 2000).
20. See, for example, Hugh Heclo, "Issue Networks and the Executive Establishment," in *The New American Political System*, ed. Anthony King (Washington, D.C.: American Enterprise Institute, 1978), 87–124. See also Paul A. Sabatier and Hank C. Jenkins-Smith, eds., *Policy Change and Learning: An Advocacy Coalition Approach* (Boulder, Colo.: Westview, 1993), 105–128.
21. John W. Kingdon, *Agendas, Alternatives, and Public Policies* (Boston: Little, Brown, 1984).
22. The term *satisficing* was coined by Herbert Simon. See, for example, Herbert Simon, *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization* (New York: Free Press, 1945).

23. Thomas A. Birkland, *After Disaster: Agenda Setting, Public Policy, and Focusing Events* (Washington, D.C.: Georgetown University Press, 1997).
24. See Alex P. Schmid and Janny de Graaf, *Violence as Communication: Insurgent Terrorism and the Western News Media* (Beverly Hills, Calif.: Sage, 1982).
25. See John Mueller, *Overblown: How Politicians and the Terrorism Industry Inflate National Security Threats, and Why We Believe Them* (New York: Free Press, 2006).
26. The Nobel laureate economist Joseph Stiglitz and Linda Blimes estimate the cost of the war in Iraq to be more than \$3 trillion. See Joseph E. Stiglitz and Linda J. Blimes, *The Three Trillion Dollar War: The True Cost of the Iraq Conflict* (New York: W. W. Norton, 2006).
27. Francis X. Clines, "Schools Await High-tech Help for Evaluating Who's a Threat," *New York Times*, October 24, 1999.
28. Dave Cullen, *Columbine* (New York: Twelve, 2009).
29. Mancur Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Cambridge, Mass.: Harvard University Press, 1965).
30. Amartya K. Sen, "Rational Fools: A Critique of the Behavioral Foundation of Economic Theory," *Philosophy and Public Affairs* 6 (1977): 317–344.
31. Howard Margolis, *Selfishness, Altruism, and Rationality: A Theory of Social Choice* (Cambridge, United Kingdom: Cambridge University Press, 1982).
32. Dipak K. Gupta, *Path to Collective Madness: A Study in Social Order and Political Pathology* (New York: Praeger, 2001).
33. Deborah A. Stone, *Policy Paradox and Political Reason* (Glenview, Ill.: Scott, Foresman, 1988).
34. For an example of factual information designed to change behavior, see D.A.R.E.'s Web site at www.dare.com.
35. For an exciting discussion on forecasting, see John L. Casti, *Searching for Certainty: What Scientists Can Know about the Future* (New York: Morrow, 1990).
36. See, for example, Peter Bachrach and Morton S. Baratz, *Power and Poverty: Theory and Practice* (New York: Oxford University Press, 1970).
37. E.E. Schattschneider, *The Semisovereign People: A Realist's View of Democracy in America* (New York: Oxford University Press, 1960).
38. Anthony Downs, "Up and Down with Ecology: The Issue-Attention Cycle," *Public Interest* 32 (Summer 1972): 38–50.
39. Richard Cohen, "We Can Never Be Really Certain of Eliminating Mistakes in Capital Cases," *Washington Post*, April 20, 2000.
40. For an excellent discussion, see Anderson, *Public Policymaking*, 3d ed.
41. In a continuing national survey, the Gallup organization found in May 2009 that 37 percent of the public approved and 57 percent disapproved of the job Congress was doing. See Gallup, "Congress and the Public," 2009, www.gallup.com/poll/1600/Congress-Public.aspx.
42. Joseph P. Kalt and Mark A. Zupan, "Capture and Ideology in the Economic Theory of Politics," *American Economic Review* 74 (1984): 279–300.
43. Edward J. Mitchell, "The Basis of Congressional Energy Policy," *Texas Law Review* 57 (1979): 591–630; James B. Kau and Paul H. Rubin, "Self-Interest, Ideology, and Logrolling in Congressional Voting," *Journal of Law and Economics* 22 (1979): 365–384; Joseph P. Kalt, *The Economics and Politics of Oil Price Regulation: Federal Policy in the Post Embargo Era* (Cambridge, United Kingdom: Cambridge University Press, 1981).
44. Daniel A. Mazmanian and Paul A. Sabatier, *Implementation and Public Policy* (Lanham, Md.: University Press of America, 1989); Sabatier and Jenkins-Smith, *Policy Change and Learning*, 105–128; Hank C. Jenkins-Smith and Paul A. Sabatier, "Evaluating the Advocacy Coalition Framework," *Journal of Public Policy* 14 (1994): 175–203; Hank C. Jenkins-Smith, Gilbert K. St. Clair, and Brian Woods, "Explaining Change in Policy Subsystems: Analysis of Coalition Stability and Defection over Time," *American Journal of Political Science* 35 (1991): 851–880.

45. Paul A. Sabatier and Hank C. Jenkins-Smith, "The Advocacy Coalition Framework: An Assessment" (paper presented to the Department of Political Science, University of Amsterdam, February 1997), 7–8.
46. Brian A. Ellison, "The Advocacy Coalition Framework and Implementation of the Endangered Species Act: A Case Study in Western Water Politics," *Policy Studies Journal* 26 (1998): 11–29.
47. Melvin J. Dubnick and Barbara A. Bardes, *Thinking about Public Policy: A Problem-solving Approach* (New York: Wiley, 1983), 203.
48. In this example I measure output in dollars. However, we need not gauge public policy outputs in monetary terms. For example, we can measure the output of a drug education policy by the size of the target audience reached.
49. Hanrahan, "Motorists Taking San Diego's Red-light Cameras to Court."
50. Ibid.
51. Michael H. Reggio, "History of the Death Penalty," in *Society's Final Solution: A History and Discussion of the Death Penalty*, ed. Laura E. Randa (Lanham, Md.: University Press of America, 1997), 1–11.
52. See Jasper Ridley, *Statesman and Saint: Cardinal Wolsey, Sir Thomas More, and the Politics of Henry VIII* (New York: Viking, 1982), 66–67.
53. Reggio, "History of the Death Penalty," 10.
54. *Gregg v. Georgia*, 428 U.S. 153 (1976); *Thompson v. Oklahoma*, 487 U.S. 815 (1988); *Penry v. Lynaugh*, 492 U.S. 302 (1989).
55. William Glaberson, "Fierce Campaigns Signal a New Era for State Courts," *New York Times*, June 5, 2000.
56. Garry D. Brewer and Peter deLeon, *The Foundations of Policy Analysis* (Homewood, Ill.: Dorsey, 1983), 386.
57. Ibid., 385.
58. Lester and Stewart, *Public Policy*, 2nd ed., 156–157.
59. Louis Jacobson, "City, State Lobbyists Work to Head Off Base Closings," *Planning* 65 (August 26–27, 1999): 26.

5

Critical Thinking and Research Design

Within a mere six-month span in 2009, four professional sports teams in New York—the Mets, Yankees, Giants, and Jets—opened the doors to brand new stadiums. By 2011, several other professional New York-area teams (including the Rangers, Liberty, Knicks, Nets, Devils, Islanders, and Red Bulls) plan to be playing in new or renovated stadiums, hoping to lure fans, corporate sponsors, and free-agent athletes.¹ These cushy, high-tech stadiums come with huge price tags. The costs of these projects often exceed \$500 million, and the cost of the legendarily spendthrift Yankees' new stadium is well over \$1 billion. Therefore, team owners nearly always ask for public financing of their extravagant sports facilities.²

Professional sports have come a long way since May 15, 1862, when William Cammeyer enclosed the ballpark on Union Grounds in Brooklyn, New York, to keep out nonpaying spectators. Today, professional sports teams—participants in a multibillion dollar enterprise—seek partnerships with cities all over the country.³ Suppose the professional sports team in your city requests public financing of a new facility. You have been hired by the city to conduct an analysis. How do you proceed?

In the course of a day, decision makers in public organizations are bombarded with questions: What are the present and future needs of our clients? Should we invest in that project? Of three possible projects, which one is best? When can we expect to finish this project? How much, how many, when, which one? As the questions pile up, decision makers can make up their minds by drawing on personal experiences or by acting on hunches. If they are really desperate, they may call on their policy analysts or go outside their organizations and hire consultants to provide them with the answers. If these individuals want to make sense of a jumble of seemingly meaningless numbers—or if they want to use reports submitted by consultants or

in-house analysts—they will need a deep appreciation for the abilities and limits of the various tools of statistics and operations research.

FIVE STEPS OF OBJECTIVE ANALYSIS

The strength of modern social science is its ability to translate vague philosophical discourse into rigorous objective analysis. The *New Webster's Dictionary of the English Language* defines objectivism as a "doctrine which postulates that reality exists independent of mind," an outlook free of prejudice, feeling, or subjectivity.⁴ This outlook, also known as objective professionalism, has shaped the twentieth-century development of Western social science in general and that of policy analysis in particular. **Objective professionalism** is rooted in inductive logic—the practice of inferring a general conclusion from premises that can be verified empirically, through observation or experience. In contrast, medieval reasoning depended to a large extent on deductive logic, in which conclusions were drawn from premises whose validity was never questioned. Despite empirical evidence supporting Galileo's discovery that the earth revolved around the sun, the scientific authorities of his day could not accept his conclusion. They rejected his finding because it contradicted the established, church-sanctioned view, which was based purely on faith.

With the birth of scientific reasoning, the authorities gradually accepted empirical verification as the best way to conduct scientific inquiry. In fact, today, empirical verification is the hallmark of objective, scientific reasoning.⁵ However, this principle runs counter to human nature, for we humans think subjectively, unscientifically. To gather information, we observe the world through filters of culture, upbringing, knowledge, values, ideologies, tastes, and personal interests. Despite our subjective perceptions, however, our reasoning is shaped by our communal existence in a scientific society, which has developed a framework of objective analysis. Our paradigm of scientific reasoning has helped us identify the following five-step process for analyzing a complex problem:

1. Define the fundamental issue and lay out the goals of the analysis.
2. Identify the alternative courses of action.
3. Forecast the consequences of the alternatives.
4. Compare and evaluate systematically all possible outcomes.
5. Choose the most preferred alternative.

We must begin our objective analysis by clearly defining our investigation's content and goals. In other words, we must make sure that we know what the fundamental issue is and what we want to establish by this analysis. In the case of stadium financing, we might ask ourselves, for whom are we building the stadium, and what needs will it serve? We also must define the goal of our study. In other words, for whom are we conducting the analysis? Although the proposed ballpark may benefit the city and downtown businesses (including the team owners), it may harm some of the other neighborhoods in the city, or it may draw resources from regional or other local governments, such as the county or a neighboring city. The project may be good for the city's sports fans but bad for those who had hoped for a new library. It may favor one professional sports team

(say, football) at the expense of another (say, baseball). Before starting our analysis, we must determine the purpose of our study.

Having defined our objective, we move to the second step of problem analysis, identifying the alternative courses of action. Our choice of alternatives will be guided by two factors: consistency and feasibility. Our alternatives should be consistent with our goal. While considering the proposed new stadium, we will want to identify the alternatives. Our first choice may be between doing the project and doing nothing. Next, we might consider building in several locations, ever mindful that we can choose not to do the project if none of the other alternatives meets our goals.

We then proceed to the third step in objective public policy analysis, forecasting the consequences of the alternatives. During public debates, the supporters of a new stadium will note the benefits of sports to the community—from creating jobs to reducing youth crime—in the loftiest terms. As an analyst, you must make an exhaustive list of these possible benefits.⁶ You may note the following reasons for providing subsidies to a private enterprise, such as a professional sports team:

- *Increase in tax revenue.* Cities prosper when more people and businesses move in. As more people work, live, and spend their income in a city, they generate revenues for the city in the form of income taxes, sales taxes, license fees, other fees, and even fines. Building a new stadium may mean that more people will work for the club, more fans will flock to the games, and more businesses will flourish because of the increased economic activity.

- *Positive externalities.* Many local governments value major league sports teams because they elevate their cities to the “major leagues.” “If the Jacksonville Jaguars aren’t well known in other parts of the country, they will be—soon,” announced the lead story in a Jacksonville newspaper on the eve of an American Football Conference (AFC) championship game against the Tennessee Titans in 2000.⁷ “Basically, this game will be a three-hour commercial for the city and the team,” said Mike May, director of communications for the Sporting Goods Manufacturers Association. Suddenly a second-tier city in Florida was talking about worldwide exposure. Some stadium proponents even suggested that the civic pride generated by a better sports facility would reduce crime among area youth.⁸ The project’s supporters also predicted increased business for the local service sector. Between the AFC championship game and the Jaguars’ matchup against the Miami Dolphins the previous week, city officials estimated that more than 50,000 visitors had traveled to Jacksonville and spent \$1 million in the city, with \$800,000 going to local hotels.

In addition to considering these benefits of the stadium project, you must evaluate the stadium’s costs. Publicly financing a privately owned sports club could have the following negative effects:

- *“Blackmail” potential.* A professional team creates a strong constituency of fans. As a result, teams demand new stadiums or expensive repairs to existing facilities. If their demands are not met, they threaten to leave their cities. In December 1982, after a highly public five-month battle leading to an agreement between the Raiders, a National Football League franchise, and the city of Los Angeles, the city controller announced

that the Los Angeles Coliseum was on the verge of bankruptcy. The controller estimated that because of the deal extracted from the city, the Coliseum would run a deficit of \$4 million in the third year of its lease.⁹ Two years after the Raiders' agreement with the city, at least thirteen of the forty-two cities with professional sports franchises were asked to provide their teams with tax incentives or to improve their sports facilities, or risk losing their teams.¹⁰

- *Wealth transfer.* When a city provides a team with subsidies, it transfers wealth from city taxpayers to wealthy athletes and multimillionaire club owners. Because the tax burden falls only on city residents, out-of-town fans receive, in effect, subsidized tickets.
- *Negative externalities.* Building a new stadium, or renovating an existing one, often displaces residents or small businesses from surrounding areas. For many of them, the cost of relocating far exceeds what the city is willing to give as compensation.

In the fourth step of objective policy analysis, we conduct a systematic *valuation of all possible outcomes*. At this stage, we place monetary values not only on the outcomes that are readily measurable in monetary terms but also on the intangible aspects of positive and negative externalities. Unfortunately, such valuations often lead to heated controversies.¹¹ In 1998, when San Diego was debating the merits of Proposition C, which would have authorized public funding of the Padres' new baseball park, city hall projected that it would need an investment of \$275 million. However, the city's consultant, Deloitte & Touche, sidestepping the fact that the project could not sustain itself from its own revenues, estimated that the project would generate \$1.8 million annually in net revenues.¹²

Finally, after conducting a thorough analysis, we *choose the most preferred alternative*, in light of the goal of the project and on the basis of some *decision criteria*. These decision criteria can take several forms. When the costs of the alternatives are the same, the criterion should be to maximize the total benefits. On the other hand, when the benefits are the same, we should choose the alternative that costs the least. When both costs and benefits vary, we try to maximize the net benefit (the difference between benefits and costs).

In real life, our criteria for thinking critically about a project can become complex, particularly when we include uncertainty in our thinking. When faced with uncertainty, decision makers may not choose the alternative with the highest expected net benefit. If you are an extra-cautious person (in technical terms, a **risk averter**), you may choose the certainty of low earnings over the lure of more attractive but uncertain returns. On the other hand, if you are a **risk taker**, you may aggressively pursue high returns, regardless of the risk. Your ability or willingness to take a risk plays a significant role in the real world—especially in a democratic, pluralistic society—because you will find that many of the steps in policy analysis do not provide certain, definitive answers. To deal with uncertain outcomes, we must follow a number of specific strategies (see chapter 13).

SETTING GOALS

Defining the goals of a project may be the most difficult step in policy analysis, because they are set not by analysts but by elected officials or higher level government officials.

TABLE 5.1 Multiple Goals: Building a New Stadium

Project goals	Expected benefits	Expected costs
Enhance the city's image	National and international exposure from having a major league sports franchise	Excessive traffic on already heavily used highways—city could be labeled overcrowded
Foster urban renewal	Extensive construction of new hotels, restaurants, and other business facilities	• Uprooting of existing businesses • Loss of identity of established communities • Relocation of low-income residents in other poor and already congested areas of the city
Generate tax revenue for the city	Increased city revenues through licenses; fees; and property, sales, and income taxes	• Tax subsidies for the sports franchise • Increased financial liabilities for the city, particularly if the team does not live up to expectations

Quite often public policy goals are not well formed, even in the minds of those who make the ultimate decisions. In inviting a professional team to a city, the mayor and other elected officials may be thinking only about the immediate benefit of creating a better image of the city. An analyst, then, should point out the other goals the city ought to consider.

The project might have multiple goals, some of which may be vague or ill defined, while others may be contradictory. A city may build a stadium to attract a new sports franchise, which will enhance the image of the city.¹³ If the city builds in a blighted area, the project may have an urban renewal objective. At times these goals may be complementary, while at others they may be contradictory. A stadium may help rebuild the neighborhood around it, but if it creates congestion and traffic problems by clogging up the surrounding highways, it may ultimately harm the city's image. In any case, when you set multiple goals, you should conduct your analysis with respect to each of them. You may present your finding in a table (see Table 5.1), so that decision makers can honestly debate the project's merits.

The primary goal of any public policy is maximizing social welfare, but we have yet to find a way to define "social welfare" to everyone's satisfaction. Terms like *the people* and *the nation* are staples of political rhetoric. Politicians of all shades, from democrats to demagogues, take liberty with these terms. The president sets a national agenda and boldly declares what the nation needs. The opposing party claims to know what "the people" really want. The question of who is "the nation" or who are "the people" has occupied some of the best minds in our intellectual history. French philosopher Jean-Jacques Rousseau, the most celebrated proponent of "general will," assumed its existence as a self-evident truth. As political theorist George Sabine points out,

The general will . . . presented a unique fact about a community, namely, that it has a collective good which is not the same thing as the private interests of its members. In some

sense it lives its own life, fulfills its own destiny, and suffers its own fate. In accordance with the analogy of an organism, which Rousseau had developed at some length in the article on political economy, it may be said to have a will of its own, the "general will (*volonté générale*)."¹⁴

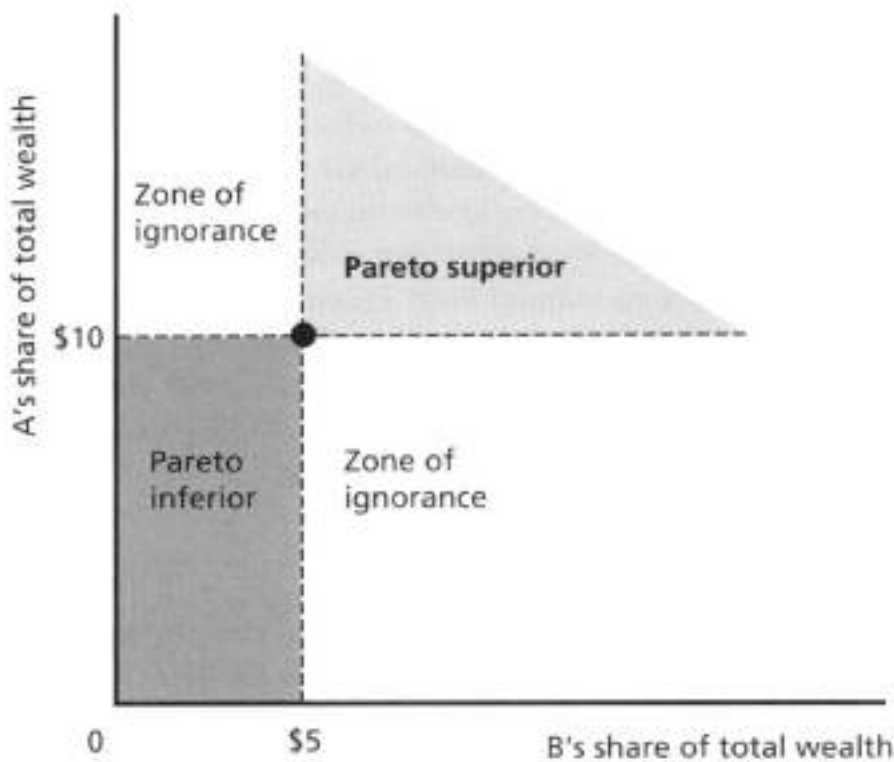
The problem with accepting the existence of an overarching **social objective** is that it often conflicts with individual freedom and individual rights. If you assume the existence of a collective will or, in economic terms, a social objective, you implicitly assume its predominance over individual aspirations. If you believe that for the world's collective good, the hunting of whales must stop, the interests of the whaling industry and the people dependent on it must be sacrificed for the greater good of the community. Similarly, public officials fund antipoverty programs by taxing the upper and middle classes. Building a freeway helps some groups (communities near the freeway, the trucking industry, motorists, the suburbs) while hurting others (communities bypassed by the freeway, the railroad industry, people and businesses displaced by the highway, the city center). Because very few public policies let everyone win and no one lose, to make wise policy choices, we must find a way of comparing the winners' gains with the losers' losses.

Utilitarians and the Pareto Principle

Toward the end of the eighteenth century, a group of British social thinkers called the **utilitarians** tried to define "social welfare." Jeremy Bentham, the best known member of this group, argued for the maximization of total utility in society. The utilitarians, in effect, argued that the marginal utility of a dollar diminishes with a person's level of income. A wealthy person, they asserted, would probably pay less attention to the loss of a dollar than would a poor person. Extending this logic, you could argue that by taking, say, \$1,000 from the wealthiest individual on Earth and giving it to a starving person, you would reduce the rich person's marginal utility to an extent less than the increase to the poor person's. As a result, the redistribution of income would increase the total utility of society. By following this formula, we could keep taxing rich people to fund social programs for poor people until the two groups' marginal utilities (one negative and one positive) were equal. The problem with such a plan is that interpersonal utilities are not measurable. We cannot quantify personal feelings about utility. If we take from a miserly rich person and give to a happy-go-lucky poor person, the redistribution of income may not affect the total utility of society at all.

Following this logic—that calculations of interpersonal utilities cannot be made—Italian economist and sociologist Vilfredo Pareto argued that a society's welfare cannot be improved by a redistribution of income from rich to poor individuals. The only way society can be considered better off is if one member gets more income (by stumbling on an oil well, for example) without taking anything away from the other members. This theory is known as the **Pareto principle** (see Figure 5.1).

Suppose our society is composed of two individuals, A and B. A commands twice as many resources as B. If as a result of a new public project A makes an extra \$10 and B makes nothing, we move along the vertical axis from the point of initial distribution between the two members. If the reverse takes place and B makes money while A makes nothing, we proceed along the horizontal axis from the initial position. Extending rules to mark A's and B's respective positions in the two scenarios, we call the shaded space

FIGURE 5.1 The Pareto Principle

bounded by these lines **Pareto superior**, the region where either A or B or both gain and neither loses. If a project places us in this zone, we should immediately recognize it as desirable. In contrast, if a project causes either individual to lose with no one gaining, we should readily discard the project, because it will take us to the **Pareto inferior** position. Public projects that benefit one group at the expense of another fall in the **zone of ignorance** and cannot be justified in the name of aggregate social welfare. The implications of the Pareto principle are extremely conservative. First, it passes no judgment on the initial distribution of wealth, regardless of how unfair that distribution may be. Second, because it is almost inconceivable that a public policy would have no redistributive component, the Pareto principle is incompatible with any sort of government activism. According to Pareto, we would quickly reject any spending on public goods—national defense, health care, access to free education, or environmental protection—because they would have to be financed by tax revenue, the burden of which would fall mostly on the upper and middle classes.

The Kaldor-Hicks Compensation Principle

Two British economists, Nicholas Kaldor and John Hicks, developed a theory of policy analysis that sought to balance the needs of society and those of individuals. Kaldor and Hicks separately argued that society was better off when pursuing projects that generated sufficiently large gains for the winners, gains great enough so that the winners could compensate the losers for their losses and still remain better off.¹⁵ Suppose a project causes A to gain \$10 while costing B \$7. A can pay B \$7 as compensation, which leaves A with \$3 of net gain. This arrangement does not violate the Pareto principle, because one

person's gain has taken nothing away from someone else. The question remains: how do we compensate the losers? In all democratic nations, if the losers can be clearly identified (as is the case when the government declares eminent domain to acquire land for building roads), the laws require that the losers be paid adequate compensation. Yet in many cases the losses are spread over a large part of society, the losers cannot be identified, or their losses cannot be adequately measured. In such cases, Kaldor and Hicks argue that we should not worry about paying compensation. As long as we can demonstrate that the benefits outweigh the costs, we can proceed with the project with a clear conscience. Such a position ultimately rejects the Pareto principle in favor of a greater "social good" but fails to help those who have no recourse when denied adequate compensation for their losses. We can cite many examples of the most well-meaning public policies' creating misery for small groups of individuals. Examples include much-needed highways that cut through the heart of a community, and dams that, while producing electricity and irrigation, flood valuable land and ruin it for inhabitants.

The Rawlsian Challenge

In contrast to the Kaldor-Hicks principle, the theory of philosopher John Rawls offers different criteria for judging social welfare.¹⁶ To deduce the fairest rule of distribution, Rawls proposes a method illustrated in the following scenario. Suppose I have a delicious pie that I plan to distribute to your class. I can cut the pie in unequal portions (with one person getting the largest share, while others literally get the crumbs), or I can divide the pie equally among the students. These portions will be distributed by a random drawing over which the students have no control. The question is, before the pie is cut, if you were given a chance to decide, how would you like it to be cut? Remember, if I cut the pie unequally, if you are lucky, you might get the biggest piece; otherwise you will surely go hungry. Rawls calls this random drawing "the veil of ignorance." According to the principles of game theory (discussed in chapter 13), when we face uncertain outcomes, the most rational course of action is to choose a strategy that minimizes our potential loss. Therefore, Rawls argues that if the students follow the precept of rationality, they should opt for an equal distribution, since that would minimize their chances of the maximum loss (known as the *minimax* strategy) of getting only a few crumbs. Extending this logic, Rawls draws a radical conclusion regarding the distribution of wealth in a society. He argues that because of the veil of ignorance (since we cannot choose our parents), if we could make a conscious decision *before we are born*, we would surely choose an equal distribution of wealth. In other words, Rawls imagined that if he could put the question to babies before they knew whether they would be born into privileged, wealthy families or to impoverished parents, they would surely choose an equal distribution. Therefore, according to the **Rawlsian criterion**, the fairness of a distribution of income should be judged by the absolute level of well-being of the community's least fortunate ones.

Suppose there are two communities, A and B, with the following sets of income distributions:

$$A = [\$10, \$25, \$8]$$

$$B = [\$200, \$150, \$5]$$

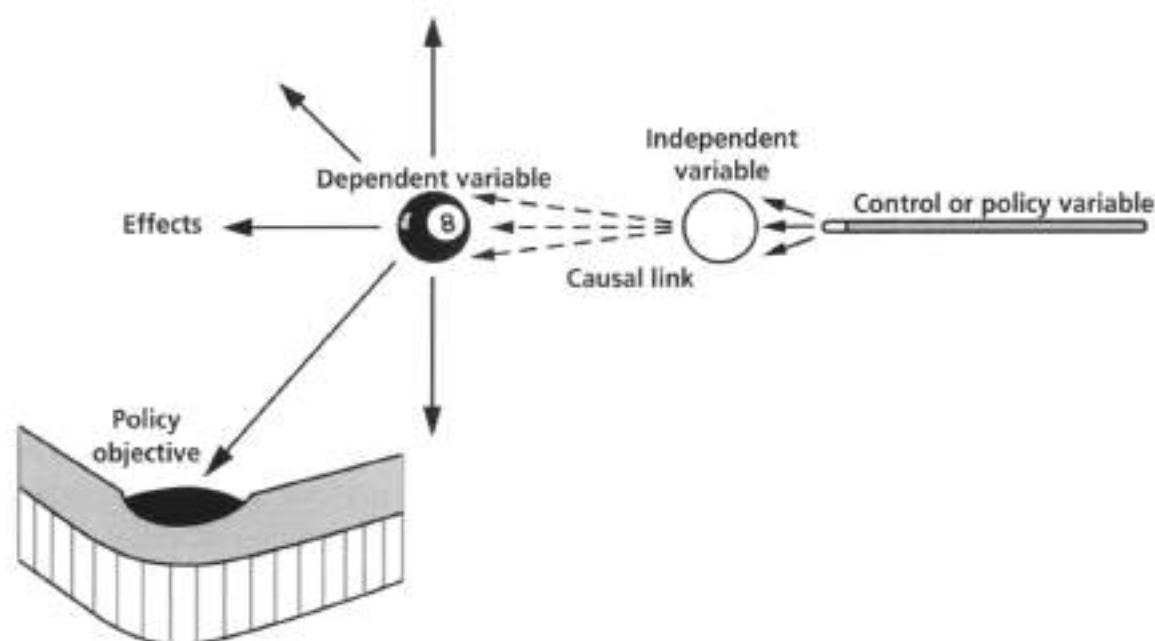
The figures in brackets show the wealth of the six members of the two communities. According to the Rawlsian criterion, you should prefer the distribution in A because the poorest member of A is better off than the poorest member of B. A drawback of the Rawlsian principle is that it throws its entire weight behind the well-being of the least advantaged members of society and disregards the welfare of others. We all know our relative positions in society, and if we are privileged members of community B, we are not likely to move to community A without serious prodding, a medicine that can be far worse than the ailment and, more important, nearly unattainable in a democratic system. However, as a standard for an ethical norm, the Rawlsian principle provides us with an extremely important method of setting public policy goals.

Setting Public Policy Goals

I may sum up our discussion as follows: When using the quantitative techniques discussed in this book, we must define a goal to be maximized. However, despite centuries of intensive searching, we have not yet found an adequate definition of social welfare. Even if we settle on one definition that benefits poor people, those who are disadvantaged, and even the largest segment of society, the goals of socialism inevitably clash with individual aspirations. And, more important, since these collective goals are articulated by a small group of powerful elite, as we have seen in communist regimes, they lead to stifled creativity, economic progress, and significant curtailment of individual rights. On the other hand, if we concern ourselves only with individual goals, we are left with extreme concentrations of income, which often ushers in political instability and social conflict. Our philosophical inadequacy and our inability to define social welfare leave a gaping hole in our understanding of project evaluation. We will return to this important problem several times in the course of this book.

CHOOSING A METHOD OF ANALYSIS

Public policy analysts strive to link project goals (output) to investment (input). The purpose of policy analysis is threefold: to choose the methods or analytical processes that will identify a **causal link** between the dependent and independent variables, to **generalize** the results, and to establish the **control**, or the policy direction. We may illustrate this complex process by drawing an analogy from the game of pool. Suppose the object of our game is to send the eight ball to a pocket by striking the white cue ball with a cue (see Figure 5.2). Our **policy objective** is to shoot the eight ball (the **dependent variable**) into the pocket by hitting it with the cue ball (the **independent variable**). The cue ball is driven by the cue, which represents the **control variable**, or the strength of our action. To properly execute our shot, we must exert correct pressure in the right direction. If the eight ball goes in the pocket, we can claim to have a successful policy. If it does not, the policy will be considered a failure. If the eight ball hits other balls on the way and makes our position on the table extremely difficult, we call this position an **unintended consequence** of the policy. Let us assume that our goal in constructing a new stadium is to increase tax revenues for the city. Our policy objective is positive net tax revenue. The dependent variable is the revenue, which is influenced by the independent variable, the new stadium. The amount of investment toward the construction is the control variable. If the new stadium causes traffic gridlock and

FIGURE 5.2 Methodology for Policy Analysis

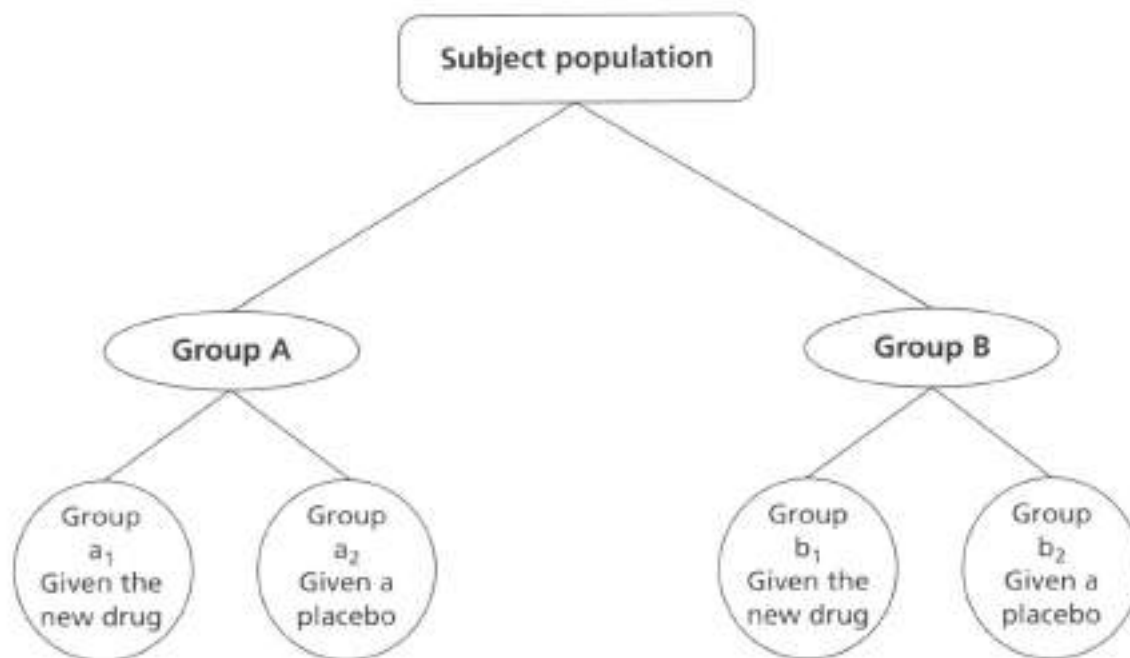
harms local businesses, we will regard these problems as the unintended consequences of the project.

Suppose we are concerned about the rising number of teen pregnancies in our state. We decide that our policy goal is to reduce the number by 15 percent in the next three years. If our analysis demonstrates a causal link between preventing pregnancy (the dependent variable) and disseminating information on reproduction through television advertising (the independent variable), we can estimate the amount of money (the policy, or control variable) we will need to accomplish our goal. In this process, then, our first task is to establish the causal relationship between the dependent and independent variables through proper **research design**.

Research design, or the plan of inquiry, links the dependent variable with a set of independent variables. It can take a number of forms that vary in scientific rigor. Broadly speaking, we classify research designs under the headings **experimental** and **quasi-experimental**. With experimental designs, we randomly assign the subjects of inquiry across experimental groups and control stimuli, and we manipulate the independent variable as we like. Using a controlled experiment, we make inferences from the results about the strength of the causal link. We then generalize our inferences to design an appropriate policy to achieve a certain goal. In contrast, with quasi-experimental research designs, we do not use these criteria. The following examples will help explain these terms.

Experimental Design

We find the classic experimental design in controlled group experiments, conducted primarily in the fields of medicine, the natural sciences, and experimental psychology. The most rigorous of all designs is the so-called **Solomon four-group design** (see Figure 5.3).¹⁷ If we want to prove the effectiveness of a new drug, we randomly assign our

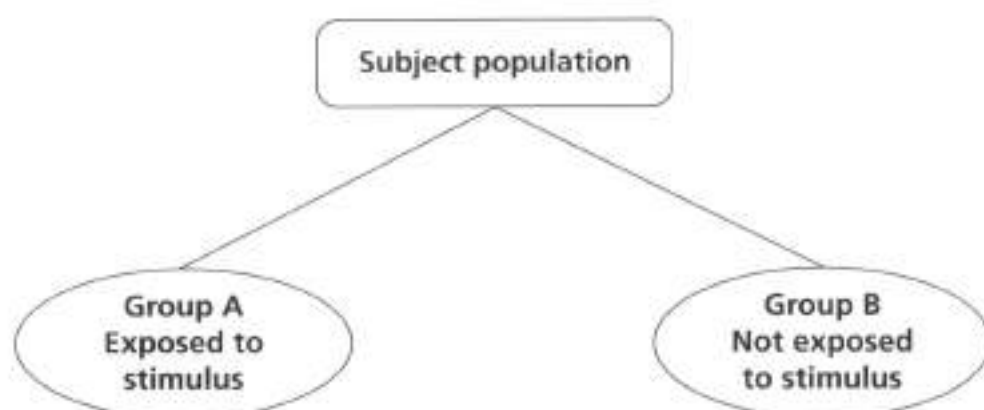
FIGURE 5.3 Solomon Four-Group Experimental Design

subjects to two groups, A and B. We further divide each group into two more groups. We give the new drug to one subject from group A and one from group B, while others receive a placebo. After administering the new drug, we monitor our patients' progress. If we find that the differences between the control groups ($a_1 - b_2$) and ($b_1 - a_2$) are significant, we can experimentally establish a causal link between the new drug and the disease.

In medical or biological experiments, researchers often exert such control over the subject population, drawing inferences and proving them with repeated experiments. New drugs come on the market, evidence mounts about various carcinogens, and clinical tests establish human behavioral patterns.

Although Solomon four-group tests are the strongest in establishing a causal link between dependent and independent variables, conducting these elaborate tests can be extremely expensive, time-consuming, and, in many cases, particularly in the area of social sciences, virtually impossible. Therefore, we must frequently be satisfied with **post-test-only control group experimental designs**, in which we simply observe the effects of experimental drugs on those who were exposed to them and on those who were not (see Figure 5.4).

In this less restricted research design, we randomly assign subjects to two, rather than four, groups: one exposed to the stimulus and one not. If we find the difference between the two groups to be statistically significant, we can establish the string of causality. Important to this research design is the random assignment of the subjects for the experiment. If we do not assign them in random order, we may allow systematic biases to creep into the experiment and contaminate the results. If we separate the groups by sex, for example, our results may be invalid if the division along gender lines introduces another independent variable along with the independent stimulus variable (in our

FIGURE 5.4 Post-test-only Control Group Design

example, the new drug). Unfortunately, however, even this less restricted research design is generally impossible to implement in the social sciences, because it requires a laboratory setup, in which subjects are randomly selected into two groups.

Quasi-experimental Design

The problem with conducting research in the social sciences is that we cannot control the real world in a manner comparable to that of a researcher working in a lab. Experimental research designs require us to distribute subjects randomly, separate them into distinct groups, guard against time lapses that dilute the effect of the stimulus, and eliminate the influence of past government policies. Therefore, in public policy research we must adopt quasi-experimental designs, in which we relax the rigors of experimental designs to meet the challenges of open society. These compromises create a number of problems for researchers.

If we want to use a strict scientific research design to test the hypothesis that the strength of family ties bears directly on a student's educational performance, we must be able to assign family stability to the subject group in a random fashion. But this clearly is not possible. Similarly, when the control variable is not amenable to change, we may have to make do with a comparison between groups that are not strictly similar. While trying to establish the relationship between family stability and educational achievement, we might find two groups, one with stability—perhaps measured by whether the parents are divorced or separated—and one without. Clearly, however, the marital status of the parents will not be the only difference between the two groups. The experiment will be affected by a variety of other intervening variables, such as the strength of the extended family, the status of religious and social community ties, and other factors.

A second problem with quasi-experimental research designs is that, in contrast to researchers in medicine, the natural sciences, and experimental psychology, those in the social sciences seldom seek to discern a stimulus-response reaction in the subject population. Instead, social scientists want primarily to understand the causal link between preexisting social or economic characteristics and revealed preferences, attitudes, values, and orientations. Suppose we try to examine the impact of a school voucher system on

a group of minority pupils. We give parents at a low-performing school the opportunity to send their children to the private school of their choice. After a year, we compare the test scores of those who went to the private schools with those who remained in the public school. If we see a significant difference between the average test scores of the two groups of children, can we positively attribute the change to the voucher system? Not necessarily—at least not according to the strictest of the scientific criteria of experimental research design. We can argue that those parents who chose to send their children to private schools may have introduced some self-selection bias into the experiment. That is, they may be more concerned about their children's education, and more involved in it, than those who did not take advantage of the voucher system. Unlike medical research, social science experiments do not permit us to randomly assign students to private schools. As a result, we cannot completely rid ourselves of the effects of self-selection and other such biases.

Another important source of bias is the passage of time. The experimental design assumes an instantaneous response to a stimulus, a response that is free of time's influence. Suppose you are conducting research in microbiology and are interested in the impact of a certain chemical compound on the growth of a particular bacterium. If after applying the compound to the bacteria-coated dish you receive instantaneous feedback, you can clearly see the test results. If, on the other hand, you must observe the gradual growth of the culture over time, you will take precautions to protect the dish from exposure to the outside environment, because you do not want the results to be contaminated by an external agent. However, in society no such protection from outside elements can be provided. This situation creates a number of intriguing problems. To continue our discussion of the voucher system, let us say that we decide to test the impact of the system on minority students by tracking those children who left the public school. If we can get data on their progress before and after the transfers, we can build a strong case for testing our hypothesis. Even under these circumstances, though, we may doubt the reliability of our data because of the possible contaminating effect of the passage of time. As children grow older, they go through many developmental stages. Critics of our experimental design may ask how much of our result may be attributed to the two school systems and how much to child development.

Take another example, national crime rates. A weary nation, jaded by sensational news stories, watched in disbelief as the crime rate declined steadily during the 1990s. President Bill Clinton claimed credit for the decrease by linking it to his recruitment of 100,000 new police officers. Matching his zeal, Republicans declared that their get-tough policies caused this welcome trend. In fact, in every city or town where the rate declined, police chiefs and city fathers claimed victory for their own programs. Can we end the debate and prove through research what reduced the crime rate? Unfortunately we cannot. To be sure, some public policies were responsible, but so were a number of broad social trends that had little or nothing to do with anticrime policies adopted at any level of government. This period saw unprecedented growth in the economy, with the lowest recorded level of unemployment in history. With more people working, fewer idle wrongdoers lived on the street. The nation was going through a major demographic shift. The children of baby boomers were passing through a crime-prone age group (between sixteen and twenty-five), but boomers had fewer children than their predecessors, causing the crime rate to decline nationwide.

You can see how difficult it is to establish a causal relationship in an open society. It is therefore not surprising that social scientists seldom resolve theoretical controversies, even when such disputes are part of accepted academic orthodoxy. So how do we discern a causal relationship between dependent and independent variables? First, we look at the strength of the correlation between the variables and, with the help of statistical tests, try to minimize the possibility of accepting a false hypothesis as true. Second, if external variables threaten to sway the results, we design the research plan carefully to remove their contaminating influence. The analyst addresses these two concerns by selecting the appropriate experimental design.

Quasi-experimental designs can take several forms: **one-shot case studies**, **correlation analyses**, **cross-sectional analyses**, or time-series analyses. Analysts typically conduct one-shot case studies on unique events of history. Consider a state road made obsolete by recent highway construction. A close-knit small town that depended on the state route is economically devastated. If you want to study the town's plight, you may need to conduct a one-shot study. The problem with one-shot case studies is that they do not give us the advantage of a valid comparison, without which we often are unable to test our hypotheses. In these studies, the best we can do is to make sure that our arguments are logically consistent and supported by valid observations. Although one-shot studies may lack scientific rigor, they can shed important light on the problem in question.

Correlation analyses consider the degree of association between two variables. If we see that an increase or decrease in one variable is associated with an increase or decrease in another, we can hypothesize a correlation between the variables. As you can see in Table 5.2, this hypothetical data set indicates that a correlation may exist between parents' marital status and children's academic achievement. We will discuss correlation analysis and hypothesis testing in chapter 7. Because researchers conduct correlation analyses by collecting comparable data for a number of cases, these analyses, unlike one-shot case studies, are amenable to statistical testing.

Analysts make cross-sectional analyses to compare the outcomes in a number of cases at a single point in time. We can compare the crime rates of similarly sized cities and ask why variations occur from city to city. The strength of cross-sectional studies is that the data are not contaminated by the passage of time. Their primary weakness concerns our ability to select cases that are directly comparable with one another. If we compare the crime rate of Billings, Montana, with that of Atlanta, Georgia, the validity of our results might be questioned because the two cities are quite different from each other. To avoid the proverbial comparison of apples and oranges, we must take special care to eliminate the effects of external variables.

TABLE 5.2**Correlation Analysis: Marital Status versus Student Achievement**

Marital status	Student achievement	
	<i>Percentage low</i>	<i>Percentage high</i>
Parents divorced or separated	60	40
Parents together	30	70

Trend analyses study a single case over time. We may conduct research on the respective crime rates of Billings and Atlanta during a certain period. Much of econometric analysis involves trend, or time-series, analysis. We will discuss time-series analysis at length in chapters 11 and 12. Its strength is that it permits us to study the same case over a long period of time. The problem with time-series analysis is that time itself has an important impact on the case we are studying. If we are studying Atlanta's crime rate over the past twenty years, we may ask whether the city has changed culturally, demographically, or economically during that period. If these variables have changed over time, they may have altered Atlanta in a significant way.

CHOOSING THE RIGHT MODEL

Because we as social scientists are restricted to the quasi-experimental design to conduct our research, we must choose a model. But what model should we use? Our fundamental problem is that we are trying to match two dissimilar worlds. We superimpose the world of quantitative techniques, based on objective, structured, scientific reasoning, on a largely unstructured, subjective world of public policy analysis. This incongruity raises two important issues: (1) A model is by definition a truncated version of reality. Can we trust it to shed light on real-world problems? (2) Models that seek to maximize social welfare assume agreement on the goal of society. How relevant are such models in a pluralistic, democratic society, in which there are multiple and diverse policy goals?

Analysts use abstract quantitative models to represent a complex reality. A model, by definition, always deviates from reality, and this deviation makes it suspect. As a model truncates reality by making assumptions, questions about its relevance become paramount. Because they use models extensively, often to reduce reality to unrecognizable forms, economists have been the objects of derision—and awe.¹⁸ We should not reject a model just because it makes simple assumptions about the real world. Because of the world's complexity, we frequently need these restrictive assumptions to separate fundamental, global patterns from local, particular, or ephemeral effects. Assumptions help us define the domain where results are expected to hold. We can then expand the boundaries of this narrow domain by systematically removing some of the assumptions to make the results resemble reality more closely. Another important reason for building models is that they help us examine the internal consistency of our reasoning. If we build a model on the premise that education discourages illegal drug abuse, we can conduct statistical tests on the basis of empirical evidence. Moreover, if we see that other variables are important in determining an individual's tendency toward drug abuse, we can take them into consideration. We can also drop a variable (for example, race or gender) if we see that it is not very useful to the explanation. Using this structured logic, models frequently explain behavior and reveal unsuspected relationships. They also eliminate the need for ad hoc reasoning and help us understand our mistakes when our predictions go awry.

Let us say that you want to forecast the number of students who might be using illegal drugs on a college campus.¹⁹ You base your forecast on a survey that shows how many freshman students used drugs last year and are continuing to do so. The survey also tells you how many nonusers have begun to use drugs since starting college. Guided by these findings, you assume that the ratios will persist throughout the college careers

of the entering freshman class. You also suppose that no external factors will have a significant effect on the trend suggested by the ratios.

We easily find reasons to question the validity of these bold assumptions. For one thing, the study does not include entering transfer students, who may behave differently than the students surveyed. In addition, the forecast does not consider the effect of education or the psychological maturity of the student population. Including these factors could alter your projections in important ways. In support of your forecasting model, you may argue that given your assumptions, you have identified the most likely trend. If you find reasons to believe that students' levels of addiction go down as they mature and become better educated, you can modify the model to allow for this change of relationship. Furthermore, if you find at the end of four years that your projections were seriously flawed, you can go back to your original set of assumptions and try to correct the mistakes. The assumptions and structure of your model allow you to process feedback in a way that would have been impossible had you made your projections based solely on "gut feelings," intuition, or ad hoc reasoning.

While it is difficult to accurately forecast the outcome of each alternative, it is even more problematic to understand and forecast the unintended consequences of government policies. These unforeseen consequences of public policies can happen because of many factors, such as a lack of scientific knowledge at the time of the adoption of the policies, an inability to fully understand complex social relationships, moral hazard, overregulation, rent seeking, and the inability to properly time a government program (see chapter 3).

FORECASTING OUTCOMES AND UNINTENDED CONSEQUENCES

The government makes decisions within the bounds of existing scientific knowledge. As science develops, it finds the follies of our previous actions. For example, in the late 1940s an agricultural research station in Mexico created a hybrid variety of rice and wheat, which was quickly described as "miraculous." It substantially increased the production of those crops and allowed many poor nations, including the perennially famine-prone India, to sustain their own food supplies. The program was so successful that it was called the "green revolution." Unfortunately, few people could have anticipated the social and political upheavals it created in India. The introduction of these new varieties of crops, which substantially increased total production, required more intensive cultivation, more irrigated water, and more fertilizer, which only the wealthier farmers could afford. The wealthier farmers also started using more mechanical devices instead of hiring day laborers.²⁰ As a result, income inequality in rural India increased, which ultimately led to social unrest; quickly, the so-called green revolution turned bloody red.²¹

Another source of unintended consequences of public policies is moral hazard (see chapter 1). It results from insulating people from the adverse consequences of their actions. For instance, terrorism expert Marc Sageman has argued that in France, extended unemployment benefits and other social safety net programs, coupled with overt and covert racism and discrimination, have made idle urban youth of North African origin into the main source of Islamic radicalism.²²

Government restrictions can create different types of unintended market distortions. For instance, in 1973, the United States learned a painful lesson about its dependence on imported oil when, as a result of U.S. support for Israel during the Yom Kippur War, the oil-producing Arab nations imposed an embargo on oil exports to the United States. Looking for alternative sources of energy at home, the nation turned toward nuclear power. However, on March 28, 1979, a valve in the Three Mile Island nuclear power generator near Harrisburg, Pennsylvania, malfunctioned and released a large quantity of radioactive coolant into the atmosphere. While the accident itself would have been a cause for concern, public anxiety was further heightened when the popular movie *The China Syndrome* depicted the nightmarish effects of a nuclear power plant accident. Suddenly, opinion swung drastically against nuclear power. Its fate was permanently sealed after the 1986 Chernobyl disaster in the Ukraine. As the government slapped stringent safety regulations on nuclear power, plant building projects were halted. As a result, the energy-hungry nation turned to the other readily available source of energy—coal. Unfortunately, coal is the worst producer of greenhouse gases, which many suspect contribute to global warming. In any case, sensing the public mood, the government's restriction of the nuclear industry gives us a clear example of market distortion from government overreaction.

Government involvement in the market can also produce the unintended consequence of creating artificial entry into the market. This results in rent-seeking behavior (see chapter 3). Firms that benefit from the restrictions suddenly capitalize on excess demand over its supply; these firms reap the benefits in terms of windfall profits (see "A Case in Point").

DESIGNING POLICY RESEARCH

Our discussion confirms that there is no easy and elegant escape from the theoretical quagmire in which we find ourselves when we try to define social welfare. As professors Edith Stokey and Richard Zeckhauser eloquently put it,

Philosophers and economists have tried for two centuries to devise unambiguous procedures for measuring and combining welfare of two or more individuals to provide a measure of total social welfare. Their quest has been as successful as the alchemists' attempt to transmute lead into gold. The occasional flickers of hope have all been extinguished; not only have no feasible procedures been developed, none are on the horizon.²³

Confronted with limited insight into the question of how to define social welfare, Stokey and Zeckhauser offer pragmatic advice. They suggest that an analyst take a limited and practical view, seeking the best solution within a narrowly defined boundary instead of the one that optimizes outcomes globally, across boundaries.

Limiting the Number of Alternatives

In our stadium financing example, our alternatives were to build the arena or not build it. Our decision to build may be complicated by the issue of location. For a project like this one, a number of sites are possible. Some may be attractive but also prohibitively expensive because of the land value. Others may offer all the desirable characteristics (adequate lot size, traffic flow, parking, and so on) but may be in environmentally



A CASE IN POINT

Halliburton and Government Patronage: Rent Seeking during the War

Stories of war profiteering have been told for as long as large armies have marched across the land. The demand for logistical supplies along with the risk of being on the front line of an active hostility creates the perfect opportunity for many to make huge profits. The case of Halliburton during the Iraq war energized the critics of President George W. Bush's policies like few others. Competitive bids for government contracts are the essence of efficient practices. They allow market forces to determine the best supplier at the least possible cost for government projects. However, war has its own logic of expediency, which often trumps concerns for efficiency. As the U.S. military invaded Iraq, the need to supply food and fuel became paramount. Since quick actions were needed, Halliburton and its subsidiary, Kellogg Brown & Root (KBR), were awarded a series of "no-bid" contracts—without having to compete with other firms. This abandonment of the normal bidding process elicited charges of foul play and corruption from many quarters. Furthermore, in the view of economists, the problem became one of rent seeking. That is, if Halliburton did not have to submit a bid to receive a contract, the government was putting up barriers to all other companies to compete with it. Therefore, Halliburton could potentially charge more than it should have by taking advantage of this barrier.²⁴

Erle Halliburton, who was born in poverty and worked in the dusty oilfields of Oklahoma and West Texas, started an oil company in the early 1920s. Today, the company is a huge conglomerate, with its declared revenues surpassing \$18 billion in 2008. Since Vice President Dick Cheney,

sensitive areas. If you want to evaluate each possible location, you may find that the cost and time required for the study are excessive. In this case, you should eliminate those alternatives that clearly are inferior to the more feasible ones and those that obviously are not workable because of cost or other factors.

Consider Only Changes in Social Welfare. Suppose you are analyzing a program to provide inner-city youths with summer jobs. Rather than getting hopelessly entangled by the question of how to stop inner-city poverty and by the larger problem of how to eliminate social injustice, you decide to evaluate the program incrementally, asking how much this program will change the youths' economic condition. In doing so, you limit the evaluation to a manageable framework by concentrating only on the marginal instead of the total change in social welfare.

Use Income as a Proxy for Welfare. "Money cannot buy happiness" goes the proverb, implying that the relationship between wealth and happiness (or utility) may not be a direct one. Yet in practice, we have no choice but to accept money as the measure of individual utility. In a situation in which one gains at the expense of another, we must be able to compare the gains and the losses in monetary terms. When we accept this proposition, we escape from the theoretical quicksand of utility measurement, only to

seen as a chief architect of the unpopular invasion of Iraq, had served as the chief operating officer of Halliburton, it quickly became the poster child of Bush administration's cronyism; Cheney, fairly or unfairly, became the face of public corruption. Although the company agreed to pay \$579 million in fines resulting from an allegation of bribing foreign government officials (not in Iraq), the largest ever imposed on a U.S. company,²⁵ Halliburton vigorously denied allegations of corruption, overcharging, and war profiteering. It argued that in supporting U.S. war efforts from military bases in Iraq and in Afghanistan, and from facilities at Guantanamo Bay, the company and its operators have taken an undue amount of risk. The substantial payment for its services by the U.S. government, Halliburton claimed, reflected that risk.

Although Cheney was most closely associated with influence peddling by Halliburton, its association with power in Washington predates the former vice president. In fact, the company developed its political ties by teaming up with a young and ambitious Democratic member of the U.S. House of Representative from Texas, Lyndon B. Johnson, over fifty years ago.²⁶ There is no doubt that Halliburton, like other companies large and small, has cultivated its political patrons. Its practices are rooted in the time-honored nexus of public need, private profit, and war making.

Discussion Points

1. How do governments create conditions for rent seeking?
2. Would you have conducted the government's operations any differently than the Bush administration, considering that the circumstances called for quick action?

land in another sort of trap—the confusion of imputing monetary values to gains and losses. Be assured, though, that our new trap is a practical problem with practical solutions rather than a theoretical problem with no solution at all.

Accepting or Rejecting a Project

To summarize, the utilitarian position does not provide us with workable criteria for analyzing public policy, because we have no way of measuring interpersonal utility. The Pareto principle turns out to be equally unhelpful, because it builds barriers against government intervention. The Kaldor-Hicks compensation principle does not address the concerns of those who might lose as a result of a public project, while the Rawlsian criterion focuses too narrowly on the welfare of the poorest members of a community. Setting aside these unworkable theories, you should make your recommendations on the basis of the following criteria.

You should accept the project if the following are true:

1. The gains are much larger than the losses, the winners and losers are roughly equal in social stature, and the losses pose only a minor problem for the losing group. Suppose you are evaluating a project to build an access road linking two major

thoroughfares. The majority of inhabitants will benefit from the road; however, a small group of people living near the project are complaining about increased traffic, with all its accompanying problems. The commissioned environmental impact report detects no significant loss to the environment or to the community. In such a case, you may recommend the project over the objections of the adversely affected group on the basis of this first criterion.

2. The benefits to society greatly exceed the cost to a specific group, or the cost of inaction is enormous. The government's power of eminent domain entitles it to purchase, confiscate, or expropriate private properties for the greater public good as long as it provides property owners with adequate compensation. In the United States we usually associate eminent domain with the government's takeover of private lands to build highways. Other countries use similar powers to confiscate property or nationalize businesses. In any case, countries justify such actions by citing the overwhelming needs of society as compared with the costs to a specific person or a small group.²⁷

3. The benefits of a proposed policy to the winners are greater than the costs to the losers, and the gains equalize some losses resulting from past discriminatory policies. Many affirmative action plans and other projects to remedy racial inequality, such as forced school desegregation and busing, would fall in this category. In addition, if a group, community, or state has received special benefits in the past, as did California during the 1980s, when the state profited from increased defense expenditures, we may make a case for bypassing it when the time comes to allocate a second round of benefits. We also may argue for equalizing the beneficial effects of public investment and diverting the second-round benefits to communities that did not receive money the first time.

In contrast, a proposed project should be rejected if the following are true:

1. The net benefit is positive, but the cost imposed on a specific group is significant. Many of our laws enforcing the Endangered Species Act of 1973 (ESA) are based squarely on this principle. Although a project may produce great benefits for the community, if it also may cause irreparable damage to a threatened species, the project will be rejected (see "A Case in Point").

2. The effect of redistribution is highly desirable, but the project does not pass the test of positive net social benefit. We frequently encounter cases in which the cause is noble but the costs do not justify action. It is desirable to provide the most comprehensive health care to everybody in the country regardless of income, age, or level of care needed. Yet not only are the costs of doing so extremely high, but any attempt to provide comprehensive government-funded health care is deemed by many as against the principles of free market capitalism. Because of their opposition, the United States, unlike most European nations and Canada, has been unable to meet the health care needs of the poorest of its citizens, although President Barack Obama has advocated passing more comprehensive health care legislation.

These criteria are by no means perfect. However, given the practical needs of public agencies, the guidelines are probably the best we can come up with. When we support our analysis with generally accepted guidelines of fairness, we make our analysis more efficient and more acceptable to the vast majority of people.



A CASE IN POINT

Do You Give a Dam? The Fate of the Snail Darter

One of the most celebrated controversies in the battle to protect endangered species involved the snail darter, a three-inch-long member of the perch family. The rare fish's habitat was threatened in the mid-1970s by the construction of Tellico Dam on the Little Tennessee River.²⁸ After four years of construction, local landowners filed a lawsuit charging that the dam violated the Endangered Species Act (ESA), bringing the \$100 million project to a halt.

This event divided the country. The United States was in a recession caused largely by a nagging energy crisis, and most studies showed that the benefits of the project exceeded the costs. The Tennessee Valley Authority (TVA) estimated the net benefits from the project to be about \$11.51 million (approximately \$33.5 million in 2000 constant dollars). Although the estimated net benefits later were revised to \$3.66 million (about \$10.65 million in 2000 dollars), the project's support among southern legislators was overwhelming.²⁹ Responding to their demands, Congress tried to bypass the ESA by exempting the Tellico project from the law. A highly reluctant President Jimmy Carter, aspiring to leave a legacy as an environmentalist, signed the exemption measure in 1979. The day after the president's action, the TVA's bulldozers started rolling. The project was quickly completed, destroying the snail darter's habitat.

Nearly a decade later, long after the dust had settled on the banks of the Little Tennessee River, policy analysts proved that the controversy had revolved around a dam that should not have been built and a species that did not require protection. They found the projected benefits to be even smaller than TVA's revised estimates. The project continued to run at a loss even before environmental impacts were factored into the calculation.³⁰ The fish thrived in the undammed parts of the river. On August 6, 1984, the U.S. Fish and Wildlife Service downgraded the status of the snail darter from "endangered" to "threatened."

Discussion Points

1. How was the Tellico Dam project stopped and how was it revived? What criteria did policymakers on both sides of the argument use to evaluate the dam?
2. What lessons can we draw from this case study about the problems of critical thinking?

Ensuring the Process Is Fair

We will now explore issues of justice and fair play from the standpoint of procedural justice. The goal of **procedural justice** is not to alter the final outcome to achieve fairness but to ensure that the process by which resources are allocated is fair. We can promote the notion of fairness by using quantitative techniques to expose hidden assumptions and values.³¹

Lotteries are unfair because one ticketholder wins most of the money, while small amounts go to other winners. For the majority of players, the lottery yields nothing. Why do we not complain about the gross inequality? We do not protest because we

know that the process of choosing the winner is fair. That is, we accept the final outcome because we believe that each person buying a ticket has an equal shot at winning the prize money. We know the process by which a winner is determined (because the drawings are often televised), and when purchasing a ticket, we agree to play by the rules of the lottery. If these conditions were violated, we would be extremely dissatisfied with the outcome. When people go over the speed limit and get a ticket, they seldom complain about the ticket itself, focusing instead on the amount of the fine or the way they were treated by the police or the court system.

Political scientist Robert Lane points out that procedural justice has four important components. It must include dignity, relief from procedural pain, a uniform standard of justice, and justice itself.³² The process must recognize, protect, and preserve the self-esteem of every individual. If you are stripped of your dignity, even the most equal distribution of wealth will seem oppressive, arbitrary, and capricious. Every individual should be assured of the swift disbursement of justice. If the process is cumbersome and time-consuming, justice cannot be served. The process also must guarantee resonance between the standards of justice followed by the judge and those recognized by the judged. If you are judged by a standard that is completely alien to you, you cannot accept the verdict as just. Finally, procedural justice must include some minimum guarantee of economic well-being.

CHALLENGES TO CRITICAL THINKING: BIASES IN REASONING

Critical thinking requires that we conduct dispassionate, objective analyses of real-life problems with uncertain outcomes. Yet our biases and preconceived notions prevent us from being completely objective in our analysis. To help us guard against biased results, we must know the most common pitfalls of critical thinking. Research in social psychology has shown that people rely on a few principles to simplify the complex task of assigning probability. Although these principles are useful in appraising an uncertain situation, they sometimes lead to systematic biases that cause severe errors in judgment.

These principles, which Amos Tversky and Daniel Kahneman call **heuristics**—the way we see and understand the world—permit us to predict the probability of rain by looking at the sky or to judge distance by looking at the size of an object (although our judgments may be flawed or inaccurate).³³ Tversky and Kahneman reduced these heuristic principles for analytical thinking into three broad categories: **representativeness**, **availability**, and **adjustment and anchoring**. When asked to make a judgment about an object or event, we make decisions on the basis of representativeness, or how closely the object or event resembles others in our experience. We also make subjective judgments on the basis of the availability of information. Finally, when our predictions differ from reality, we use new information to adjust and anchor our appraisal of the future.

Representativeness

At the root of all scientific discovery is association. If we draw a connection between the coat of a woolly caterpillar and the coming winter, or between a groundhog's shadow and the imminence of spring, we have in effect predicted the arrival of the seasons. Such associations have led scientists to make discoveries by causally linking events. At the

same time, reasoning by association also has produced old folk tales and prejudices. When we cringe at the sight of rowdy youths in a subway but look favorably on individuals wearing business suits, we are making a judgment on the basis of representativeness. Because we often base our associations on long-term observations, our predictions often turn out to be correct. However, unless we are careful, blind adherence to association can lead us to biased policy decisions.

Guard Against Biases of Irrelevant Information. Biases may result from our insensitivity to the probability of outcomes. In a psychological experiment, subjects were told that a group contained 70 percent farmers and 30 percent lawyers.³⁴ If we choose an individual at random from this group, the probability that we will select a farmer is 0.7. The subjects in this experiment used this prior probability to compute the odds of getting a farmer or a lawyer in a random pick. Subjects were given worthless information on a person from the group, such as

Dick is a thirty-year-old man. He is married with no children. A man of high ability and high motivation, he promises to be quite successful in his field. He is well liked by his colleagues.

After reading this passage, subjects were asked to estimate the probability that Dick was a farmer. Because the passage contains no information about Dick's profession, subjects should have stuck to the overall probability of 70 percent, the proportion of farmers in the group. Yet they ignored the prior probability and assigned a 50 percent likelihood of Dick's being from either group. In life we often are misled when confronted with useless information, which can trigger judgments on the basis of representativeness. Despite conflicting prior information, we decide to ignore it.

Pay Attention to the Sample Size. The second set of problems may arise from our insensitivity to sample size. Suppose I ask you the following question:

In your city, you can have your baby delivered at one of two hospitals, one large and one small. The probability that you will have a boy is 50 percent, and the chance that you will have a girl is also 50 percent. If each hospital records the days in a year when the percentage of boys exceeds 60 percent of its births, which hospital can expect to have counted more such days at the end of the year?

You would be incorrect to choose the large hospital or to assume that both hospitals would have an equal number of days with 60 percent or more boys. This is because, as we will see later in the book, as the number of tries increases, the samples become less likely to stray from the average probability figure. Therefore, large hospitals, with more births, are less likely than small hospitals to record days with more than 60 percent male births. If you made the wrong choice, take heart in the fact that many people make decisions on the basis of faulty statistical reasoning—including the people whose profession is to teach statistics.

Remember that Chances Do Not Self-Correct—They Merely Dilute. Suppose you are losing a game of cards. Do you tell yourself that because you have had a string of bad luck, your luck is due for a change? If so, you are making an erroneous inference about an independent probability distribution. As you may recall from our discussion

of the independence of probability, if you receive four heads in a row in a coin toss, your chances of getting tails do not go up on the fifth try. Thinking that chances self-correct is a common mistake. But the anomalies of chances are not corrected over many tries; they are merely diluted. This idea is especially true for small samples. Remember that the smaller the sample, the less representative it will be of the larger population. Therefore, we place less faith in results based on small samples than in those drawn from large ones.

If You Do Not Have the Relevant Information, Do Not Predict. Suppose your department wants to complete a project in a tight time frame. You have been asked to recommend one out of numerous applications from a group of vendors. Because the vendors are not aware of the time constraint when they apply for the contract, their applications do not mention their ability to deliver on time. Instead, the applications contain the usual information: lists of projects the vendors have been involved in, the qualifications of the project team, and, of course, budget estimates.

What would you do in this situation? You might look into vendors' relative experience and qualifications and predict their ability to meet the accelerated deadline. If you did, you would be committing the error of judgment that Kahneman and Tversky³⁵ call the error of insensitivity to predictability. Because the available information tells you practically nothing about the vendors' ability to complete the job in a hurry, and because this ability depends on their present workload and capacity to devote key personnel to the project, you would need more information to make your selection. Without this information, you would be better off making no prediction at all.

Beware of Picking the "Right" Evidence. The heuristics of representativeness require that we match an unfamiliar, observed event with a familiar one whose outcomes are known to us. Because an event has a number of characteristics, you must choose the most significant ones in determining this representativeness. Chances are good that you will choose evidence to suit your preconceived biases. This likelihood raises tough issues, especially if it involves your deeply held values and biases.³⁶

Do Not Ignore the Regression Toward the Mean. Suppose you are an avid but average golfer, and you know your handicap. This morning you drive the ball well from the tee, and you nearly get a hole in one. However, you bogey the next hole, failing to put the ball in the cup from a close distance. You are disappointed, and you have every reason to be. But statistically speaking, you should not despair, because your performance shows simply that you cannot beat the mean. In other words, you are merely playing your average game. If you make one exceptional shot after another, then you are playing at a level much higher than your natural average. English scientist Sir Francis Galton noted more than 100 years ago that outcomes tend to gravitate to the mean. He called this tendency the regression toward the mean.

Ignoring the regression toward the mean is another important source of error in subjective judgment. If you can find a long history of past performance, you will be much better off basing your prediction on the mean of that performance than on an optimistic or pessimistic extreme that reflects minor variations in external conditions.

Consider the Allais Paradox. The precepts of objective reasoning state that when faced with an uncertain outcome, you should choose the option that maximizes your expected

returns. We define expected return as the probability of winning multiplied by its reward. If I offer you \$10 for calling a coin toss correctly, your expected return is $0.5 \times \$10 = \5 . When facing two options, either the possibility of winning \$10 in a coin toss or that of gaining \$25 by rolling a die, you should stick with the coin, because the expected return of tossing the coin (\$5) is greater than that of rolling the die ($1/6 \times \$25 = \4.17). Unless you are a real gambler (a risk lover), the laws of probability dictate that despite the chance of a larger reward with the die, you should choose the coin. In real life most of us fail to follow this principle, which is known as the **Allais paradox**, named after French mathematical economist Maurice Allais.³⁷

Suppose you have been offered two options. The first option gives you a 100 percent chance (certainty) of winning \$1 million; the second option gives you a 50 percent chance of winning \$5 million. In experiments most people choose the first option in violation of the rules of expected returns ($1.0 \times \$1 \text{ million} = \1 million , which is less than $0.5 \times \$5 \text{ million} = \2.5 million). Why do people ignore the laws of probability? When we expect an outcome to occur with certainty or near certainty, we tend to focus less on the probability factor than we do when anticipating an uncertain result.

Consider two other options. In the first option you have a 1 in 10,000 chance of winning \$1 million, and in the second, your chance of winning \$20 million is 1 in 1 million. Which option do you choose? Most people choose the chance to win \$20 million, despite the fact that expected returns are higher for the \$1 million option. Again, contrary to the principles of statistical reasoning, when probabilities are small (1 in 10,000 vs. 1 in 1 million), people tend to focus more on the reward than on the odds of winning. Allais unearthed some important biases in human reasoning that often prove critical in the decision making of a public organization.

Availability of Information

We often estimate the probability of an event by observing the frequency of similar events. If your city is considering whether to build its own garbage recycling plant, you can appraise the probability that the plant will succeed by examining the available information on similar ventures. Tversky and Kahneman call this decision-making device *availability*. They assert that information on similar events yields valuable clues about future possibility. This information points toward a more certain outcome when similar examples are numerous than when they are scarce. As we gather information, the following biases may cloud our judgment: **relative retrievability**, the **effectiveness of a search**, **biases of imaginability**, and **illusory correlation**.

Biases of Relative Retrievability. Our memory works like a filing cabinet. If we know where to search, we can retrieve a stored item quickly. However, if we file it in the wrong place, we cannot retrieve it easily, even information that is quite important. One way to keep track of information is to associate an event with some recognizable pattern or some other event. We then retrieve the information we need by recalling this other pattern or event. For example, most people remember the day of the 9/11 attacks vividly. We may not recall something important that happened in the past three weeks, but we remember many details about an important national event that took place years ago. When we recall memorable events or those associated with other facts that made a strong impression on us, we are susceptible to biases in judgment. These biases may lead to biases of availability, causing us to give relatively unimportant factors more weight

than they deserve. Our awareness of life-threatening diseases may be only cursory until we hear that a celebrity has a particular malady, or until we learn about the illness of a friend or relative. Then our awareness of that disease goes up, prompting fear and, often, hysteria. We may focus on the probability of getting that particular disease more than is warranted by our personal habits, lifestyle, or genetic history. We also may disregard another source of illness, one more plausible than that suggested by the condition of the celebrity or friend.

The Effectiveness of Search. Have you ever conducted a computer search for information through a library information retrieval system? Suppose you are looking for information on a certain subject (say, drug abuse in North Carolina). If you define your key words properly for this search, you will be rewarded with information on a number of highly relevant publications. However, if your choice of key words is too wide (simply *drugs*), you will be inundated with irrelevant information. On the other hand, if you define your key words too narrowly, you may miss a large number of important works on the subject. In our minds we use key words to retrieve information effectively. Our effectiveness depends very much on our ability to define the parameters of our search.

Biases of Imaginability. If you are considering a risky venture, you would start your assessment by first listing both the difficulties and the advantages of the project. You would consider a full slate of possibilities before taking any action. Your options are limited only by your imagination. If you have difficulty estimating the relative risks and payoffs of a project, you may leave yourself open to biases. These sorts of biases seem to be most prevalent when a decision must be made about extraordinary events (such as when a disaster-preparedness program must respond to catastrophic and unpredictable events like floods, tornadoes, earthquakes, or large-scale riots).

Illusory Correlation and Organizational Myths. People often make decisions on the basis of long-held beliefs born out of illusory, or faulty, correlation. For example, do you believe that you have a "lucky" article of clothing that helps you during an uncertain situation? If so, you may be guilty of illusory correlation, which is described by the Latin phrase "post hoc, ergo propter hoc," or "after this, therefore because of this." In other words, you are taking past co-occurrences as signs of correlation. If you conclude that your decisive victory in a tennis match was due to the shirt you were wearing, you are guilty of finding a causal link between the two events (wearing the shirt and winning the match) when, in fact, there is no connection.

The errors in judgment caused by illusory correlation are often deep-rooted and cannot be easily corrected. We create myths on the basis of such errors, myths that affect not only our individual decisions but also the decisions we make at the organizational level. We find in some of the taped conversations between President Richard Nixon and his aides during the Watergate crisis classic examples of organizational decisions made on the basis of a paranoid worldview. Time and again, Nixon and his staff made decisions rooted in a fortress mentality, ultimately contributing to the demise of his administration. In any organization the vision of a decision maker can be clouded by myths of illusory correlation created by a single key staff member or by collective myths produced by a number of people in an organization.

Adjustment and Anchoring

In estimating the uncertain outcome of a project, we are frequently influenced by the first estimate. In social psychology this tendency is known as anchoring. You can witness anchoring on a popular television game show, when a group of contestants is asked to judge the price of an item. You will notice that the first contestant's answer seems to have a great deal of influence on the answers of the rest of the contestants. Contestants appear to be calibrating their own answers by the previous answers. In real life we often arrive at a quantitative judgment by working from an initial number. This tendency can cause serious errors in assessing an event's probability.

In this section I have tried to show some of the major sources of cognitive bias that can produce errors in critical thinking. The list of sources is long but by no means exhaustive. Thanks to the prolific work of social psychologists, we are learning more about these natural biases. The most important conclusion we can draw from this discussion is that these distortions of judgment are not necessarily caused by self-serving motivations such as wishful thinking, desire for reward, or fear of punishment. Instead, they are rooted in our thinking and the ways we process information. Proper knowledge of our cognitive biases may alert us to the pitfalls of objective analyses.

A FEW PARTING SUGGESTIONS

Quantitative models and numerical analyses are simply tools, aids to your natural analytical capabilities. Like all other tools, by themselves they are neither good nor bad—their ultimate value to you, your organization, and society at large depends on the way you use them. Along with knowing what you know, it is imperative to know what you do not know. Analysts should identify their zones of ignorance. By using an objective analysis, you can proceed up to a certain point, beyond which you must make judgments using personal, political, or social values. Quantitative techniques will take you to this point. To go further, you must seek help beyond the scope of these techniques. You must enter the arena of the decision makers and goal setters.

Another important contribution of quantitative techniques is that they force us to disclose our biases, hidden values, prejudices, and presuppositions. The light of quantification significantly illuminates the dark, hidden crevasses of subjective analysis. When we fully specify a problem in numerical terms, we can agree or disagree about the nature of the problem on level ground, without hyperbolic rhetoric masking our value judgments. When a policy based on an objective analysis fails, we can learn from its failure and recalibrate the policy by listening and responding to feedback.

As we have seen, the usefulness of objective techniques goes far beyond their substantive contributions to policy analysis. They also affect procedural justice in a public policy debate. Objective analysis not only provides policymakers with clearly defined evidence and a set of corresponding arguments but also gives them an intellectual framework for open discussion. Even when we find its conclusions unacceptable or infeasible, objective analysis shapes the structure, language, and issues of a policy debate.³⁸ The debate, focused on the ways of reaching the conclusion rather than on the conclusion itself, sets the stage for a broader understanding and an acceptance of the policy process. Once a policy has been adopted, its success depends on its economic viability, administrative feasibility, and, above all, political acceptability, as filtered through the legitimacy of the policymakers.

KEY WORDS

- Adjustment and anchoring (p. 122)
- Allais paradox (p. 125)
- Availability (p. 122)
- Biases of imaginability (p. 125)
- Causal link (p. 109)
- Control (p. 109)
- Control variable (p. 109)
- Correlation analyses (p. 114)
- Cross-sectional analyses (p. 114)
- Dependent variable (p. 109)
- Effectiveness of a search (p. 125)
- Experimental research designs (p. 110)
- Generalize (p. 109)
- Heuristics (p. 122)
- Illusory correlation (p. 125)
- Independent variable (p. 109)
- Objective professionalism (p. 102)
- One-shot case studies (p. 114)
- Pareto inferior (p. 107)
- Pareto principle (p. 106)
- Pareto superior (p. 107)
- Policy objective (p. 109)
- Post-test-only control group experimental designs (p. 111)
- Procedural justice (p. 121)
- Quasi-experimental research designs (p. 110)
- Rawlsian criterion (p. 108)
- Relative retrievability (p. 125)
- Representativeness (p. 122)
- Research design (p. 110)
- Risk averter (p. 104)
- Risk taker (p. 104)
- Social objective (p. 106)
- Solomon four-group design (p. 110)
- Trend analyses (p. 115)
- Unintended consequence (p. 109)
- Utilitarians (p. 106)
- Zone of ignorance (p. 107)

EXERCISES

1. Your city is considering upgrading its sewage treatment plant, and you have been hired to make a recommendation. Think critically and discuss how you would approach the problem.
2. Write an essay on the pitfalls of objective analysis. How can you reduce their effect on your reasoning?
3. Suppose your local government is proposing the construction of a halfway house for nonviolent juvenile offenders in a city neighborhood. You have been hired to evaluate the project. Explain how you would begin your job and what problems you would face in arriving at an objective conclusion.
4. Explain the Rawlsian theory of distributive justice. After reading the Government Accountability Office report on poverty in America in Appendix A, do you think the Rawlsian theory is a politically and economically viable principle to guide our national public policy?

Notes

1. Richard Sandomir, "Tickets for New Stadiums: Prices, and Outrage, Escalate," *New York Times*, August 25, 2008.
2. Mandy Rafool, *Playing the Stadium Game: Financing Professional Sports Facilities in the '90s* (Denver, Colo.: National Conference of State Legislatures, 1997).

3. Dean V. Baim, *The Sports Stadium as a Municipal Investment* (Westport, Conn.: Greenwood, 1994). See also Neil deMause and Joanna Cagan, *Field of Schemes: How the Great Stadium Swindle Turns Public Money into Private Profit* (Monroe, Maine: Common Courage, 1998).
4. Sidney R. Bergquist, ed., *New Webster's Dictionary of the English Language* (Melrose Park, Ill.: Delair, 1981), 653.
5. Karl R. Popper, *The Poverty of Historicism* (Boston: Beacon, 1960).
6. See, for example, Robert A. Baade and Richard F. Dye, "An Analysis of the Economic Rationale for Public Subsidization of Sports Stadiums," *Annals of Regional Science* 22 (July 1988): 37–47.
7. Mya M. Borger, "Having a Ball: Sunday's AFC Championship Game a Financial Jag-ernaut for Team, City," *The Business Journal* (Jacksonville, Fla.), January 24, 2000.
8. Baim, *The Sports Stadium*, 5.
9. "Danger of Coliseum Going Broke Seen," *Los Angeles Times*, December 17, 1982.
10. Arthur Johnson, "Municipal Administration and the Sports Franchise Relocation Issue," *Public Administration Review* 6 (November/December 1983): 519–529.
11. We will discuss the problems of valuation later in the book. At this point, we will keep in mind that we are always assigning monetary values to intangible items, such as a human life, a person's reputation, and the pain and suffering claimed by a plaintiff in legal proceedings.
12. Philip J. LaVelle, "Proposition C Is about a lot More Than a Ballpark," *San Diego Union-Tribune*, October 20, 1998. Needless to say, this figure of \$1.8 million in net benefits was used liberally by supporters of the project. The proposition passed by a wide margin, and by September 2000 the half-completed project had already run into financial problems.
13. In 1957 C. Norris Poulson, the mayor of Los Angeles, chided the opponents of his plan to bring baseball's Brooklyn Dodgers to the city. He charged that they wanted to keep the city in the "bush league." See "LA Council Votes Dodgers Deal, 11-3," *Los Angeles Times*, September 17, 1957 (quoted in Baim, *The Sports Stadium*, 5).
14. George H. Sabine, *A History of Political Theory*, 3rd ed. (New York: Holt, Rinehart & Winston, 1961), 588–589.
15. Nicholas Kaldor, "Welfare Propositions of Economics and Interpersonal Comparison of Utility," *Economic Journal* 39 (September 1939): 549–552. See also John R. Hicks, "The Foundations of Welfare Economics," *Economic Journal* 49 (December 1939): 696–712.
16. John Rawls, *A Theory of Justice* (Cambridge, Mass.: Belknap, 1971).
17. For a detailed discussion of research design, see David Nachmias and Chava Nachmias, *Research Methods in the Social Sciences* (New York: St. Martin's, 1976). See also Carl V. Patton and David S. Sawicki, *Basic Methods of Policy Analysis and Planning* (Englewood Cliffs, N.J.: Prentice Hall, 1986).
18. You may have heard the story of an economist who is stranded with a group on a desolate island, with only a can of beans. The economist suggests, "Let us assume a can opener."
19. For information on how to use quantitative methods for this kind of forecasting, see the discussion of Markov's chain in chapter 12.
20. For a more complete picture, see, Stanley A. Freed, *Green Revolution: Agricultural and Social Change in a North Indian Village* (New York: American Museum of Natural History, 2002).
21. See, for example, Bernhard Glaeser, *The Green Revolution Revisited: Critique and Alternatives* (London: Allen & Unwin, 1987).
22. Marc Sageman, *Leaderless Jihad: Terror Networks in the Twenty-first Century* (Philadelphia: University of Pennsylvania Press, 2008).
23. Edith Stokey and Richard Zeckhauser, *A Primer for Policy Analysis* (New York: W.W. Norton, 1978), 276.
24. See Pratap Chatterjee, *Halliburton's Army: How a Well-Connected Texas Oil Company Revolutionized the Way America Makes War* (New York: Nation Books, 2009).

25. Zachary A. Goldfarb, "Halliburton, KBR Settle Bribery Allegations," *Washington Post*, February 12, 2009.
26. Dan Birody, *The Halliburton Agenda: The Politics of Oil and Money* (New York: Wiley, 2005).
27. For a discussion of eminent domain, see Ralph C. Chandler and Jack C. Plano, *The Public Administration Dictionary*, 2nd ed. (Santa Barbara, Calif.: ABC-CLIO, 1988), 62–63.
28. Mark Van Putten and R. J. Smith, "At Issue: Has the Endangered Species Act Been a Success?" *CQ Researcher*, September 15, 2000.
29. Robert K. Davis, "Lessons in Politics and Economics from the Snail Darter," in *Environmental Resources and Applied Welfare Economics: Essays in Honor of John V. Krutilla*, ed. V. Kerry Smith (Washington, D.C.: Resources for the Future, 1988), 211–236.
30. John V. Krutilla and Anthony C. Fisher, *The Economics of Natural Environments: Studies in the Valuation of Commodity and Amenity Resources*, rev. ed. (Washington, D.C.: Resources for the Future, 1985).
31. For a classic exposition of essentially the same arguments, see Allen Schick, "System Politics and Systems Budgeting," *Public Administration Review* 29 (March–April 1969): 139–150. The flip side of these views is that an agency might use quantitative techniques either to rationalize its preordained decisions or to conduct a public relations campaign to placate critics and ward off outside interference. A later study by Harvey Sapolsky claimed that, contrary to the highly touted success stories, Program Evaluation and Review Technique network analysis had nothing to do with the progress and problems of the navy's Polaris project. See Harvey M. Sapolsky, *The Polaris System Development: Bureaucratic and Programmatic Success in Government* (Cambridge, Mass.: Harvard University Press, 1972).
32. Ibid.
33. Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science* 185 (September 1974): 1124–1131. Derived from the Greek word *heuriskein* (to find out), "heuristics" is defined by the *New Webster's Dictionary of the English Language* as "teaching principles which allow students to make their own discoveries."
34. Ibid.
35. Ibid.
36. I will illustrate my point about self-selection bias by telling a story about two friends. One friend held an unshakable faith in astrology, while the other did not. All their lives they argued about the validity of astrology. One day, the believer in astrology learned that he would have to make a long trip across the state by car. Having consulted all the relevant astrological signs, he chose the most auspicious day to start his journey, only to have a serious accident within half a mile of his home. The poor man was taken to a nearby hospital, where his nonbelieving friend went to greet him with his best-rehearsed arguments against putting faith in such superstitious nonsense. Before the visitor could say anything, the injured man looked up from his hospital bed and asked, "You believe in astrology now, don't you? Can anybody survive such an accident without guidance from the stars? I am lucky to have consulted my chart before I started my trip. What do you think of astrology now?"
37. Maurice Allais, "Fondements d'une théorie de choix comportant un risque et critique des postulats et axiomes de l'école Américaine" ("The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and the Axioms of the American School"), in *Expected Utility Hypothesis and the Allais Paradox: Contemporary Discussions of Decisions under Uncertainty with Allais' Rejoinder*, ed. Maurice Allais and Ole Hagen (Dordrecht, the Netherlands: D. Reidel, 1952).
38. Giandomenico Majone, *Evidence, Argument, and Persuasion in the Policy Process* (New Haven, Conn.: Yale University Press, 1989). See also E. S. Quade, *Analysis for Public Decisions*, 3rd ed. (New York: North-Holland, 1989).

6

Basic Statistics

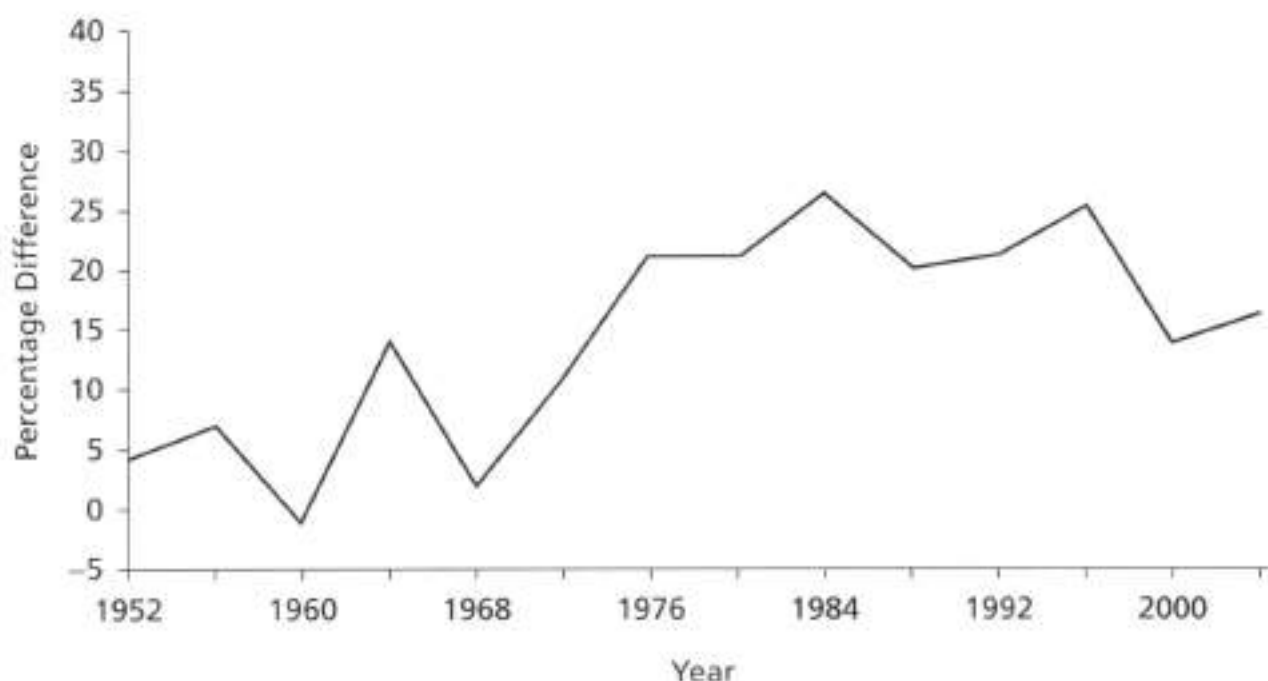
NUMBERS AS STORYTELLERS

The footprints of history are preserved in recorded information. Properly collected and analyzed numbers can shed a great deal of light on the past. We know that we can numerically extend information collected from autopsies of Egyptian mummies to the general population of the time. But numbers can even reveal seemingly nonnumerical aspects of human lives, such as feelings. For example, some political scientists claim that income class has increasingly polarized American politics.¹ The data that I have plotted in Figure 6.1 support this claim by showing the rise in differences in presidential voting choices between the classes. Now, on the basis of this scenario, you can develop many different stories about current political conditions in the United States and draw conclusions about its future. This is the task of quantitative analysis: to make sense of the story behind the numbers. Facing a torrent of information, an analyst or researcher discards the unnecessary information and arranges and rearranges the rest to create something new—a logically coherent set of arguments.

METHODS OF DESCRIPTIVE STATISTICS

There are many ways of looking at the world of information. We may look at the world **ex post facto** and **ex ante facto**—that is, what has already taken place (*ex post*) and what will take place (*ex ante*). The basic tools for analyzing *ex post* data are the measures of **central tendency** and **dispersion**, and the tool for analyzing *ex ante* data is probability theory. Certainty is what has already taken place; uncertainty is what may follow.

Analysis of data expressed in numbers starts with the measures of central tendency and dispersion. Together these two are called descriptive statistics. The term *descriptive* obviously implies description—yet from description, one starts analysis; from the understanding of the

FIGURE 6.1 Class Divisions in Presidential Voting, 1952–2004

Source: American National Election Studies, "ANES Cumulative Data File" (October 31, 2005 version, 1948–2004), www.electionstudies.org/study pages/download/datacenter_all.htm.

past, one looks for clues to forecast the future. Therefore, in this chapter we will start with the methods of descriptive statistics and correlation, and in the following chapter we will introduce the theories of probability.

Measures of Central Tendency

The aim of descriptive statistics is to describe a situation or assess the prevailing condition with numbers. In our daily communications we can describe a situation verbally without the use of numbers. Anthropologists tell of primitive tribes whose abilities to express quantities with numbers are limited.² However, the power of description increases with the proper quantification of the situation. Thus, one may describe the day's temperature as "unusually hot," or one can express the temperature as a measure of degrees Fahrenheit or Celsius. The use of a recognizable index for measuring temperature immediately facilitates its understanding at the absolute level ("it is really hot today"), as well as its comparative evaluation ("the highest recorded temperature in twenty years" or "it was hotter here today than it was in Phoenix").

A series of numbers do not readily convey a coherent picture. The purpose of descriptive statistics is to look for the number around which a series has a tendency to cluster, so that it can be seen as the representative of that series. The tendency of a numerical series to cluster around a number is called the *central tendency*. By this measure, we offer the number that best represents the series. For example, consider Table 6.1, which presents the prices of five houses in a neighborhood. To convey a quick impression about this neighborhood, we can reproduce the entire table, which is rather cumbersome, or we may try to describe it using a single number. Of the many measures

of central tendency, three are most commonly used in the area of social sciences: the mean, median, and mode. The **mean** is the arithmetic average of a series, expressed as

$$\mu = \frac{\sum_{i=1}^n (x_i)}{n} \quad (6.1)$$

where x is a variable with n number of observations, and μ is the mean.

This is the formula for the arithmetic average. Those of you who don't often look at mathematical expressions may recall that the Greek capital letter sigma (Σ) is used as a sign for summation. The term x with subscript i is a **variable** (in Table 6.1, it is the prices of the houses, which vary with each house). The subscript i refers to the specific observation, or in this case an individual house. If $i = 2$, then x_2 refers to the second observation in the series, which in this case is the price of the second house, \$65,000. The term n is a **constant** and measures the number of observations, which in this case is 5. The subscript below and superscript above the summation sign Σ are read as "the sum of variable x , with observations (i) varying from 1 through 5." Therefore, in this case, we add up the values of our five observations to arrive at \$900,000. Dividing by 5, we get the average value of a neighborhood house: \$180,000.

You may notice that I have expressed the mean with the Greek letter μ (pronounced "meu"). However, sometimes you will find the mean written as $\bar{X}$ (said as "X bar"). It is important to note that it is a common tradition in statistics to denote the mean of a **population** (the entire group in question) as μ and the mean of a **sample** (a small fraction of the population chosen by a researcher to observe) as $\bar{X}$. Unless I note otherwise, I am speaking of a population.

The arithmetic mean is the most commonly used measure of central tendency. You may notice that most quantitative techniques (many of which are discussed in this book) are built around the analytical anchor provided by the mean. This is because the implication of the mean is rather intuitive. Also, in mathematical statistics, the mean as the measure of central tendency has the highly desirable property of **asymptotic convergence**. Suppose I have an unbiased coin, which I flip, and every time I flip heads, I note it. If the coin is not flawed, as I repeat my experiment a number of times, the ratio of the number of heads to the total number of tosses will come closer and closer to .5 as the number of tosses increases. This number (.5) is the same as the theoretically derived ratio of **relative frequency**—the number of desired alternatives divided by the total number of alternatives. In this case we have just one desired alternative, heads, and two possible alternatives, heads and tails. Using this rule we can calculate that the odds of flipping heads are $1:2 = .5$. Because a die has six sides, the possibility of getting either a

TABLE 6.1 **Distribution of Housing Prices**

House	Price
I	\$65,000
II	\$65,000
III	\$150,000
IV	\$230,000
V	\$390,000

1 or a 6 with a throw of the die is $2:6 = .333$. This neat mathematical property allows the mean to be used in statistical theorems.

However, extreme values in a series influence the mean, especially when the sample size is small. In our example of the unequal prices of a small number of houses in a neighborhood, the mean price of housing is \$180,000, which is higher than two-thirds of the houses in our sample. The house priced at \$390,000 has influenced the mean.

The **median** is the middle number in a series. In our example, the median price of a house in the neighborhood is \$150,000. Because it is the middle number, it is impervious to the extremes in the series. However, determining the median requires the physical inspection of the series, and thus it cannot be calculated as easily as the mean. Also, it does not have the asymptotic property of unbiasedness that makes the mean such an attractive measure of central tendency in statistics and mathematics.

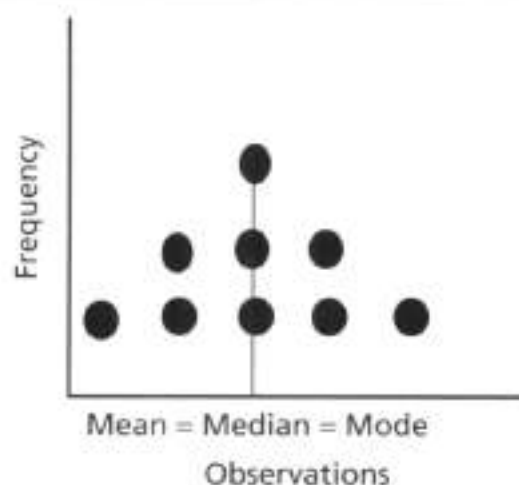
Finally, the **mode** is the most frequent number in a series. In this simple example, it happens to be \$65,000.

Since the mean price of housing is \$180,000, the median is \$150,000, and the mode is \$65,000, which of these three considerably different numbers should represent the average price of a house in the neighborhood? Let us elaborate.

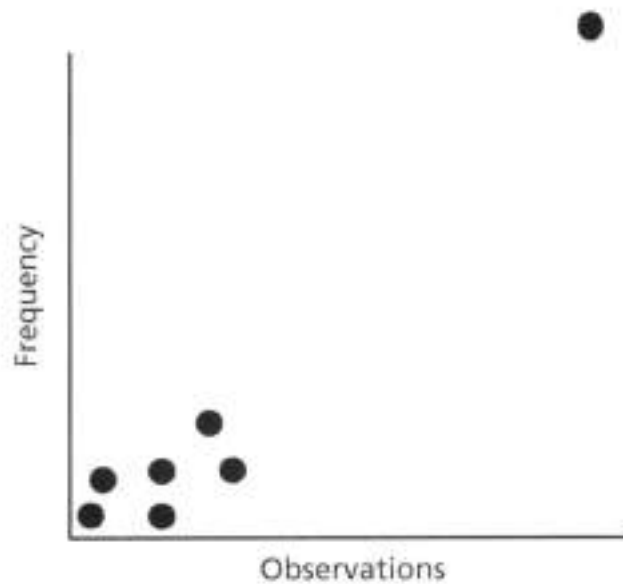
Consider, for example, Figure 6.2. The distribution is clearly symmetric. The choice of a measure of central tendency is not going to be controversial because, for this distribution, the mean, median, and mode will coincide. Controversy will soon ensue if the distribution is not symmetric. In Figure 6.3, for example, all the numbers cluster close to one another, except for one, which is an extreme value. This unusually large number will unduly influence the mean. In this situation either the median or the mode can be used to represent the series. In Figure 6.4, one observation with a particular value predominates. Therefore, the modal value will best represent the series. The extreme value makes the mean an inappropriate measure. Similarly, because of the prevalence of one value, even the use of the median may be considered less than appropriate.

FIGURE 6.2

A Symmetric Distribution



Note: The distribution of observations shows a clear tendency to cluster around the middle number. Therefore, since this series of numbers is evenly distributed, we may accept any of the three measures of central tendency as appropriate. Still, one should always calculate the mean when reporting the central tendency, unless there is a special reason not to do so.

FIGURE 6.3 Distribution with an Extreme Value

Note: When there is an extreme value, the total picture of a distribution becomes complex. In these cases, the use of any one measure of central tendency would be controversial, since one can always question the "representativeness" of the number.

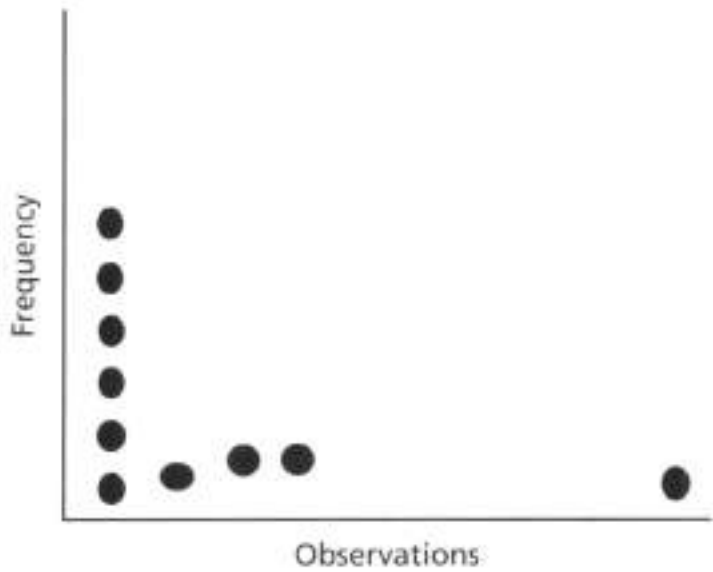
At the outset, presenting the day's temperature in your city on a local television station's weather map looks like a simple enough task. However, controversy would arise if the city recorded a high temperature at noon and the temperature then fell precipitously as a cold weather front moved in. The question would be to decide which measure to choose as representative of the day's temperature. Varied geographical areas with several small climate zones within the city would further complicate the problem. In this case, a weather station located in a coastal area may record a significantly different temperature from one located inland.

The fundamental question is, Why do such problems arise? We encounter these kinds of controversies any time we try to describe a complex, multidimensional phenomenon with a one-dimensional measure. The more complex the situation, the more problematic its representation with just one set of numbers. The alternative is to describe a multidimensional phenomenon with several different sets of numbers. However, more than one number tends to numb one's senses. Representing a complex situation with just one number may have its disadvantages, but the alternative does not seem all that attractive either. The choice of the best measure of central tendency often calls for subjective judgment.

Weighted Estimates. If I ask you, "How long would it take you to go to the airport?" you might say, "If the traffic is *exceptionally* heavy, it can take fifty-five minutes; if it is *unusually* light, fifteen minutes; but *in general*, it takes me about twenty minutes." Not knowing the traffic conditions, we may estimate the travel time to the airport by averaging the three numbers. If we do that, we come up with an estimate of thirty minutes. Notice that this time is far too pessimistic, because it is one and a half times the time it usually takes to drive to the airport. This average reflects the extreme nature of the most pessimistic assessment of the situation. In such cases, when you do not want to be influenced by the presence of extreme values, yet you do not want to disregard them, you

FIGURE 6.4

Distribution with the Modal Value



Note: This figure shows a hypothetical situation in which the use of the mode would be appropriate. Since there is an extreme value, the mean would be unduly influenced by it.

may derive your estimate by giving **weights** to the observation. That is, you may give a weight of 4 to the average time. We can write the series as shown in Table 6.2.

From these data, we can calculate the weighted average as $[(15 \times 1) + (20 \times 4) + (55 \times 1)]/6 = 25$. Notice that now we are not dividing the sum by the number of observations; instead, we are dividing the sum by the total of the weights. Also notice that this number (25) is a lot closer to the “most likely” number (20) than the unweighted average (30). By attributing weights, we have made the estimate a lot more realistic. In mathematical symbols, the calculation of weighted average is written as

$$\mu = \frac{\sum(f_i \times x_i)}{\sum f_i} \tag{6.1a}$$

where f_i represents the respective weights for the observations.

“Grouped” Data. Similar logic is applicable when data are available in **groups**. The range of a group is called the **interval**. The data on housing prices for a neighborhood may be available only in groups. I have written the housing price data given in Table 6.1 in group form in Table 6.3.

TABLE 6.2

Weighted Observations

Assessment	Time (in minutes)	Weight
Optimistic	15	1
Most likely	20	4
Pessimistic	55	1

TABLE 6.3 **Grouped Data**

Price range	Number of houses (f_i)	Midpoint of intervals (x_i)	Midpoint $\times$ number of houses ($f_i \times x_i$)
\$0–\$99,999	5	\$50,000	250,000
\$100,000–\$199,999	17	\$150,000	2,550,000
\$200,000–\$400,000	3	\$300,000	900,000
	$\Sigma f_i = 25$		$\Sigma f_i x_i = 3,700,000$

Using the formula given by equation 6.1a, we can calculate the weighted mean as $\$3,700,000/25 = \$148,000$.

The median for interval data is calculated somewhat differently. To calculate the median, first inspect the data to locate the middle observation. As you can see from Table 6.3, there are twenty-five houses in the sample. Therefore, the midpoint is located at the thirteenth house. However, we don't know the exact price of the thirteenth house in the sample series. Hence, we assume that all the observations are evenly spaced within the range, which in this case is equal to \$99,999. We can estimate the spacing of the prices by dividing the length of the range by the number of units within it. In other words, we assume that the difference between two successive houses is $\$99,999/17 = \$5,882.30$.

If we assume that within the range \$100,000 to \$199,999, the seventeen houses are evenly spaced in price, the thirteenth house (the sixth one within the price range) will be priced at $\$100,000 + (\$5,882.30 \times 6) = \$135,293.80$.

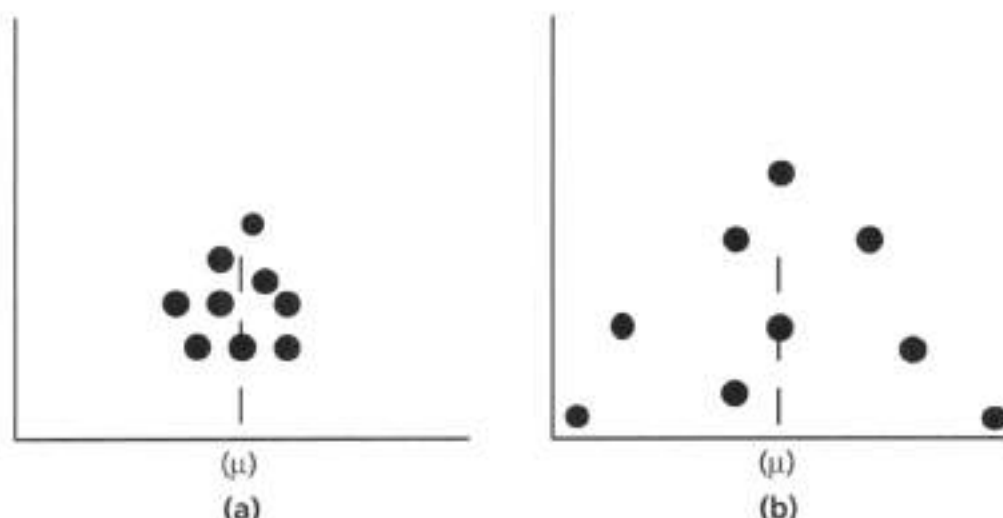
Therefore, *weighted averages are only approximations of the actual ungrouped data. So, whenever possible, use ungrouped data for the calculation of central tendency.*

Measures of Dispersion

A series is characterized not only by its central tendency but also by how strong this central tendency is. That is, how closely does the series cluster around the measure of central tendency? In Figure 6.5, both series have the same mean, but obviously the first series (Figure 6.5 (a)) is more bunched together than the second (Figure 6.5 (b)).

The closer this clustering is, the more confident we can be of the representativeness of our measure of central tendency. To understand the measures of dispersion, let us compare our first neighborhood with a second neighborhood. We can measure the relative dispersion, or the "scatteredness," of a series by using range, mean absolute deviation, variance, or standard deviation (see Table 6.4).

The **range** simply measures the difference between the highest and the lowest values in a series. In Table 6.4, the houses located in neighborhood A have a range of \$325,000, while the range for neighborhood B is \$50,000. The concept of range is also associated with what are known as the hinge and the midspread. One-quarter of the observations in the series lie below the **lower hinge**, and one-quarter lie above the **upper hinge**. The distance between the two hinges, containing half the observations around the median, is called the **midspread**. For example, the series of numbers 3, 17, 69, 85, 97, 117, 198, 211, 217, 300, and 301 does not tell us very much about the nature of its distribution,

FIGURE 6.5**The Effects of Different Dispersions**

Note: If the observations are closely bunched together, the distribution has a smaller level of dispersion (a) compared with the one in which the observations are widely scattered (b).

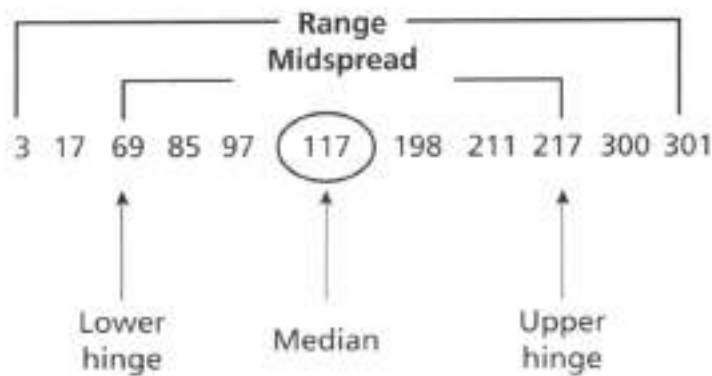
but we may express it with the help of the range, the median, and hinges, as shown in Figure 6.6. This depiction sheds more light on the characteristic of the series. As you can see, in this series, 117 is the median value of the series of eleven numbers, the third and the ninth observations are the two hinges, and the difference between them is the midspread.

To measure how scattered or dispersed a distribution is, we need to find out the deviation from the mean, or the average distance between an individual observation within the series and the mean. However, if we add up the different distances from the mean, the positive numbers cancel out the negative numbers. Thus, we get the results shown in Table 6.5 if we subtract the mean value of the series (\$180,000) from the various values of the housing prices of the two neighborhoods. From this table, it should be apparent that if we try to compare the deviation from the mean for the two series, the sum of column 3 ($x_i - \mu$) will give us 0, as will the sum of column 6 ($x_j - \mu$). Therefore, it is important to note that *the sum of deviations from the mean is always equal to 0*.

We can avoid the problem of measuring dispersion from the mean by either disregarding the signs of the deviations or by squaring them (since the square of a negative

TABLE 6.4**Relative Distribution of Housing Prices between Two Neighborhoods**

House	Neighborhood A	Neighborhood B
I	\$65,000	\$150,000
II	\$65,000	\$180,000
III	\$150,000	\$180,000
IV	\$230,000	\$190,000
V	\$390,000	\$200,000

FIGURE 6.6 The Range, the Midspread, and Hinges

number is positive). The first method of calculating the average dispersion of a series is called the **mean absolute deviation**, written as

$$\text{Mean absolute deviation} = \frac{\sum_{i=1}^n |x_i - \mu|}{n} \quad (6.2)$$

Using this formula, we can see in Table 6.5 that the mean absolute deviation for the first neighborhood is \$104,000 = (\$520,000/5), and for the second, \$12,000 = (\$60,000/5). According to the mean absolute deviation measure, the second neighborhood is nearly 8.7 times more homogeneous (less dispersed) in its distribution of housing prices than the first neighborhood.

However, the mean absolute deviation poses two problems. First, computationally it is cumbersome; second, it does not possess some of the most desirable mathematical properties. Therefore, we may try another way to eliminate the negative signs in the deviation from the mean—that is, by squaring the deviations. This process can be written as

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (6.3)$$

This is the measure of **variance**, expressed as σ^2 . To calculate the variance of the two neighborhoods, we need to square the deviations from the mean (the third and sixth

TABLE 6.5 Relative Distribution of Housing Prices: Deviation from the Mean (in dollars)

House	Neighborhood A			Neighborhood B		
	(x_i)	$x_i - \mu$	$ x_i - \mu $	(x_i)	$x_i - \mu$	$ x_i - \mu $
I	65,000	-115,000	115,000	150,000	-30,000	30,000
II	65,000	-115,000	115,000	180,000	0	0
III	150,000	-30,000	30,000	180,000	0	0
IV	230,000	50,000	50,000	190,000	10,000	10,000
V	390,000	210,000	210,000	200,000	20,000	20,000
Total		0	520,000		0	60,000

TABLE 6.6 Calculation of Variance

House	Neighborhood A (squared deviation from the mean) $(x_i - \mu)^2$	Neighborhood B (squared deviation from the mean) $(x_j - \mu)^2$
I	13,225,000,000	900,000,000
II	13,225,000,000	0
III	900,000,000	0
IV	2,500,000,000	100,000,000
V	44,100,000,000	200,000,000
Total	73,950,000,000	1,200,000,000

columns in Table 6.5), add them up, and divide by 5. The problem with variance is that the numbers are often extremely large (see Table 6.6). In this case the calculated variances for the two neighborhoods are 14,790,000,000 ($73,950,000,000/5$) and 240,000,000 ($1,200,000,000/5$). We can reduce the size of the variance by taking its square root. The resulting number is called the **standard deviation** of the series. Because of its several highly desirable mathematical properties, the standard deviation (expressed as σ) is by far the most frequently used measure of dispersion. The standard deviation of a series is calculated using the following formula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (6.4)$$

By taking the square roots of the variances of the two series, we calculate the standard deviations, which turn out to be \$121,614.14 and \$15,491.93. Obviously, these numbers are much more manageable than the variances. From these measures, we can see that the dispersion for the first neighborhood is more than seven times the dispersion for the second neighborhood.

When we are speaking in terms of population variance and standard deviation, we use the Greek alphabet: σ^2 and σ . When we consider sample variance and standard deviation, we use the Roman (English) alphabet: s^2 and s . This is a rather important convention in statistical terminology: *When denoting population mean, variance, and standard deviation, we use the Greek alphabet: μ , σ^2 , and σ . However, sample mean, variance, and standard deviation are written $\bar{X}$, s^2 , and s .*

At this point, you should note another significant difference between the formula for calculating the population standard deviation σ and sample standard deviation s . The sample standard deviation, which is only an estimate of the population standard deviation, must be corrected for bias by dividing the sum of squares by $n - 1$, instead of n . This implies that as the size of the sample (n) increases, the subtraction of 1 will have less and less impact on the calculated value of the standard deviation. I discuss the rationale for subtracting 1 from the number of observations, known as the **degrees of freedom**, in the next chapter. The formula for calculating the **sample standard deviation** is given by

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1}} \quad (6.4a)$$

TABLE 6.7 Comparing the Dispersion of Two Distributions

Series I	Series II
0	0
1	5
2	10
3	15
4	20

Dispersion of Distributions. Suppose we want to know which of two distributions has more variations within it. Unfortunately, unlike the mean, we cannot readily use standard deviations to compare the relative dispersion of distributions. The problem with the standard deviation is that it is an absolute measure. That is, it is influenced by the unit of measurement. Consider two sample distributions, I and II. You will notice that series II is series I $\times 5$ (see Table 6.7).

The sample standard deviation (s) is calculated to be 1.58 for series I and 7.91 for II. Because series II is five times the value of series I, so is the value of the respective standard deviation. This can sometimes pose a practical problem. Suppose we want to compare the dispersion of housing prices in two diverse cities, Tijuana, Mexico, and San Diego, California. These two cities probably offer the greatest contrast in relative prosperity among neighboring cities around the world. When we try to compare the two, we face a problem: Because housing prices in San Diego are much higher than those in Tijuana, the dispersion in the San Diego real estate price will be magnified by the absolute difference in the price level. However, it is entirely possible that the disparity in housing prices is greater in Tijuana than in San Diego. To compare the relative dispersion between the two cities, we must divide the standard deviation by the mean. This measure, known as the **coefficient of variation**, is written as

$$\text{Coefficient of variation} = \frac{\sigma}{\mu} \quad (6.5)$$

Looking at the two neighborhoods in Table 6.4, we can calculate the coefficient of variation for them as

$$\text{Neighborhood A} = \frac{135,968.70}{180,000} = 0.755; \text{Neighborhood B} = \frac{18,708.30}{180,000} = 0.104$$

The calculated coefficients of variation show that the variance for the first neighborhood is more than 7 times (7.26, to be precise) that of the second one. Incidentally, you may also notice that by using this measure, we conclude that despite the difference in scale, the distributions of values for series I and II have the same relative dispersion.

Skewness and Symmetry of Distribution. A question often asked about a series of numbers is, Where are the mean, median, and mode of the distribution? The positions of these three measures of central tendency can give decision makers extremely valuable clues about the distribution of the series. Take, for example, the series of numbers shown in Table 6.8.

The data from Table 6.8 are plotted in Figure 6.7. The perfect bell-shaped distribution shown in this figure is called symmetric. For distributions like these, all three

TABLE 6.8

Data Showing a Symmetric Distribution

Observation	Frequency
1	5
2	8
3	12
4	15
5	21
6	35
7	21
8	15
9	12
10	8
11	5

measures of central tendency—the mean, median, and mode—fall on the same point. If you calculate the mean for the distribution shown in Table 6.8, you can see that it is approximately (with the appropriate rounding off for the decimal point) equal to 6. Also, since there are five numbers above and five numbers below, the median for the distribution is also 6. Finally, because the observation 6 has the highest frequency, it is also the modal point for the distribution.

However, this happy situation changes if the distribution is tilted to the right or to the left of the perfect bell shape. A tilted distribution is called a **skewed distribution**. A distribution tilted to the right is called a **positively skewed distribution**; one that tilts to the left is called a **negatively skewed distribution** (see Figure 6.8). Consider, for example, two new series along with the one given in Table 6.8. Data for three different series are presented in Table 6.9. The symmetric series is called frequency A, and the two subsequent series are labeled frequencies B and C. If you plot these three series, you will see the curves shown in Figure 6.8.

Notice that the median for the three distributions remains the same. Because we have eleven observations, the sixth observation is the middle point for each of these

FIGURE 6.7

Example of a Symmetric Distribution

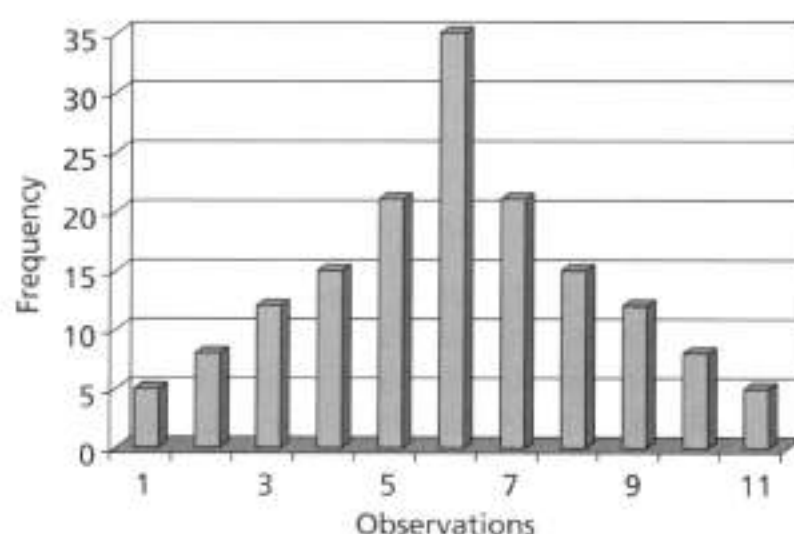
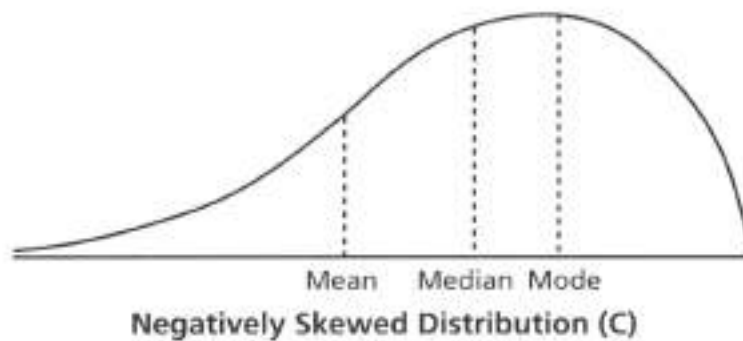
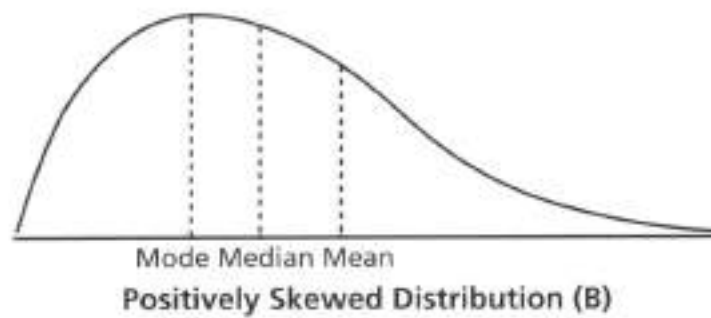
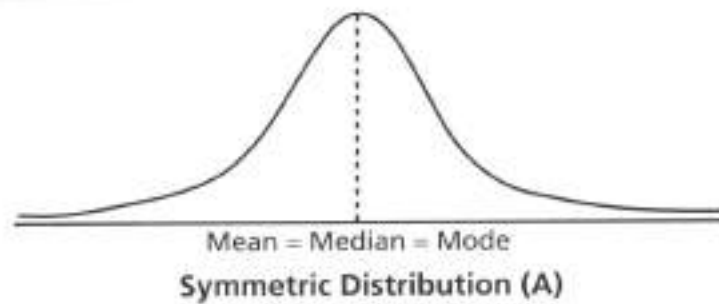


FIGURE 6.8**Comparison of Symmetric and Asymmetric Series****TABLE 6.9****Examples of Symmetric and Asymmetric Distributions**

Observation	Frequency A	Frequency B	Frequency C
1	5	25	5
2	8	35	6
3	12	20	7
4	15	18	8
5	21	15	10
6	35	12	12
7	21	10	15
8	15	8	18
9	12	7	20
10	8	6	35
11	5	5	20

TABLE 6.10**Mean, Median, and Mode in Symmetric and Asymmetric Series**

	Mean	Median	Mode
Symmetric distribution (A)	6.0	6.0	6.0
Negatively skewed distribution (B)	4.3	6.0	2
Positively skewed distribution (C)	7.6	6.0	10

distributions. The means for series B and C are 4.3 and 7.6, respectively. The mode for series B is 2; and for series C it is 10. Table 6.10 summarizes this information.

For a policymaker who needs to make a quick judgment about a data series, it is extremely important to remember that *for a negatively skewed distribution, the mean is less than the median, and for a positively skewed distribution, the mean is greater than the median*. These two figures can give a rough picture of the distribution. In fact, you may notice that the greater the difference between the two, the greater the extent of skewness of the distribution.

Which Measure of Central Tendency to Use

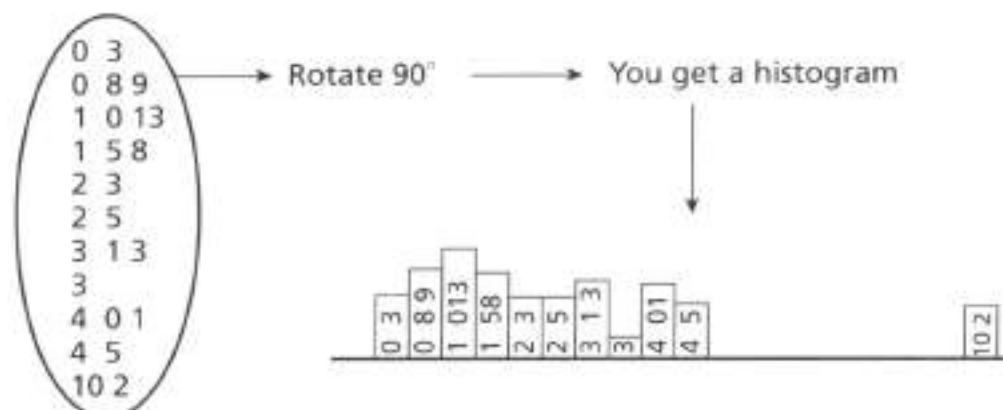
Previously I argued that the choice of a measure of central tendency depends on the subjective assessment of the person making a statement. However, there are a few rules of thumb:

1. In a symmetric distribution, all three measures are the same. Therefore, one could choose any one of the three.
2. In a highly skewed distribution, the median may be the most useful measure of central tendency.
3. An overwhelming preponderance of one number may indicate the use of the mode.
4. The arithmetic mean should be used to calculate the average, unless there is a special need to do otherwise.
5. If for any reason the median or mode is used, this should be made clear to the reader.

A Quick Glance at the Distribution: The Stem-Leaf Method

A series of numbers often does not convey a coherent picture. For example, assume that over a fifteen-week period, local police officers arrested varying numbers of individuals on drug-related charges: 15, 23, 8, 31, 9, 45, 41, 18, 11, 3, 13, 25, 33, 40, 10, 102. Clearly, one cannot draw many conclusions from these numbers. Nevertheless, a policy analyst must frequently draw quick conclusions about a distribution, such as housing prices in a neighborhood, the ages of children in a detention center, or the time taken by various employees to complete a task. In such cases, the **stem-leaf method** offers a helping hand.

The stem-leaf method arranges a series of numbers by *stems* and *leaves*. For example, in the preceding series of numbers, the first digit of a number is the stem, and the second digit is the leaf. For the number 15, 1 is the stem and 5 is the leaf. For a single-digit number, such as 3, the stem is 0 and the leaf is 3. You may notice that you can produce a histogram or a bar chart simply by rotating the stem-leaf diagram (see Figure 6.9).

FIGURE 6.9 Stem-Leaf Method and Histogram

The advantage of the stem-leaf method is that it presents a lot more information regarding the distribution. For instance, we can tell from Table 6.11 that although the data vary a good deal, the median value, 25, is much closer to the minimum, 3, than to the maximum, 102. Therefore, it is clearly a negatively skewed distribution. The output also indicates that the maximum value is an outlier, or a real exception to the rest of the values in the series. In many decision-making instances, this method can provide some rather important insights.

Correlation Coefficient: Pearson's r

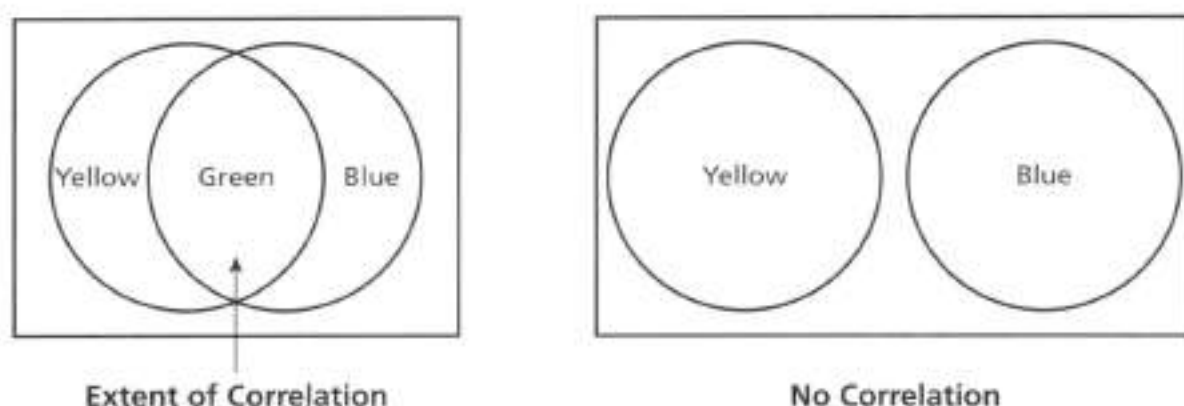
You may often wonder if two sets of data are correlated. Throughout the ages, patient observers have gained scientific knowledge by discovering close associations between two phenomena. Ancient astronomers noted that the tides change with changes in the lunar cycle. In social science research, the extent and nature of the association between two variables are often the subject of inquiry.

We may want to know if being a victim of child abuse is correlated with later criminal behavior. The nature of the correlation between two variables is described as positive or negative. In Figure 6.10 two variables, Yellow and Blue, are shown with Venn diagrams. If the two overlap, there is a correlation. As the two move toward each other, the

TABLE 6.11 Stem-Leaf Analysis: Drug Arrests, $N = 16$

0 3	
0 8 9	
1 0 13	Lower Hinge
1 5 8	
2 3	
2 5	Median Value
3 1 3	
3	Upper Hinge
4 0 1	
4 5	
Outside value	
10 2	

Note: Minimum = 3.000; lower hinge = 10.500; median = 25.000; upper hinge = 36.500; maximum = 102.000.

FIGURE 6.10 The Extent of Correlation between Two Variables

extent of correlation increases. The two circles correlate perfectly when they completely overlap and thus become indistinguishable from each other. If they do not correlate at all, they do not intersect.

If an increase in one variable corresponds to an increase in the other, there is a **positive correlation**. If the presence of childhood abuse is linked with higher levels of criminal behavior, we call it a positive correlation. In contrast, the presence of an abusive relationship linked to a child's lowered academic achievement is an example of a **negative correlation**. I have shown this in Figure 6.11.

We can numerically measure the extent and strength of correlation with the help of Pearson's r , or simply the **correlation coefficient**. Suppose we have two sets of variables, X and Y . The correlation coefficient between them is measured by the formula

$$\text{Correlation coefficient} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}} \quad (6.6)$$

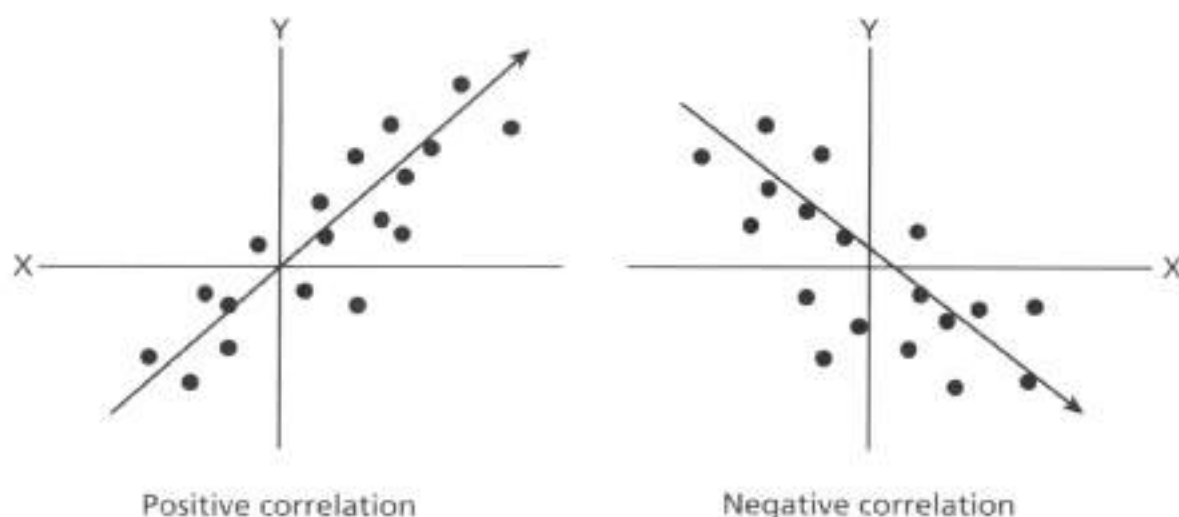
FIGURE 6.11 The Nature of Correlation

TABLE 6.12 **Crime Rate and City Size**

City	Population ($\times 1,000$)	Crime rate (per 1,000 inhabitants)
A	50	1.5
B	57	1.7
C	65	1.6
D	68	2.9
E	88	2.0
F	92	3.1
G	96	2.5
H	100	3.5
I	120	3.8
J	125	3.7

The size of the coefficient indicates the extent of correlation. If it is equal to 1, there is a perfect correlation (or a complete overlap) between the two variables, and if it is equal to 0, the two variables are independent of each other (there is no overlap). The sign of the correlation coefficient shows the nature of their relationship. If the sign is positive, the relationship is positive, and vice versa. Therefore, the correlation coefficient equals 1 when there is a perfect positive correlation, it equals -1 when there is a perfect negative correlation, and it equals 0 when there is no correlation at all.

Let us suppose the crime rate in your city has gone up in recent years. Of late, a few local newspaper articles have condemned this trend. Your police chief, however, contends that because the city is growing rapidly in size, the crime rate can be expected to rise as well. You want to know if there is a correlation between the crime rate and the population size. You have collected the data series shown in Table 6.12. Using the formula in equation 6.6, you can calculate the correlation coefficient, which turns out to be .8655. The positive number corroborates your police chief's assertion regarding a strong correlation between the crime rate and the population size of a city.

KEY WORDS

Asymptotic convergence (p. 133)

Central tendency (p. 131)

Coefficient of variation (p. 141)

Constant (p. 133)

Correlation coefficient (Pearson's r) (p. 146)

Degrees of freedom (p. 140)

Dispersion (p. 131)

Ex ante facto (p. 131)

Ex post facto (p. 131)

Grouped data (p. 136)

Interval (p. 136)

Lower hinge (p. 137)

Mean (p. 133)

Mean absolute deviation (p. 139)

Median (p. 134)

Midspread (p. 137)

Mode (p. 134)

Negative correlation (p. 146)

Negative skew (p. 142)

Population (p. 133)

Positive correlation (p. 146)

Positive skew (p. 142)

Range (p. 137)

Relative frequency (p. 133)

Sample (p. 133)

Sample standard deviation (p. 140)

Skewed distribution (p. 142)

Standard deviation (p. 140)

Stem-leaf method (p. 144)

Upper hinge (p. 137)

Variable (p. 133)

Variance (p. 139)

Weight (p. 136)

EXERCISES

1. What is the implication of the name *descriptive statistics*? What are the measures of central tendency and dispersion? Explain the term *central tendency*. Why do we face controversies regarding the appropriate choice of a measure of central tendency?
2. Consider the following series and choose the most appropriate measure of central tendency for each series. Explain your choices.

Series A	Series B	Series C	Series D
10	5	3	5
11	8	7	7
9	11	9	9
27	13	25	11
16	16	7	9
15	18	7	7
17	21	7	5

Suggestion: You may want to plot the data to determine the patterns.

3. What do we measure by the coefficient of variation? Give an example of its possible use in public sector decision making. Using this method, comment on the relative dispersions of the four series of data presented in the preceding table.
4. Which one of the measures of central tendency is most commonly used in mathematics and statistics? Why?
5. How is the correlation coefficient useful? Give an example of its possible use and abuse.
6. We often hear the argument that increased size of the government retards economic growth. The Central Intelligence Agency collects the most reliable cross-national data in its *World Factbook*. Go to its Web site at <https://www.cia.gov/library/publications/the-world-factbook/index.html> and collect data on the size of government (budget revenue/gross domestic product) for Japan, Australia, and New Zealand, and for Western Europe and North America. Also gather data on rates of growth with regard to per capita income. Prepare a table, calculate the correlation coefficient between the two variables, and explain your finding.

Notes

1. Jeffrey M. Stonecash and Mack D. Mariani, "Republican Gains in the House in the 1994 Elections: Class Polarization in American Politics," *Political Science Quarterly* 115 (Spring 2000): 93–113.
2. The people of the Nambiquara tribe of the Mato Grosso forest in Brazil, for example, lack any system of numbers. The closest they come to expressing equality between two sets of items is by using a verb that means "they are alike," *The Guinness Book of Records 1992* (New York: Facts on File, 1991), 269.

7

Probability and Hypothesis Testing

As the old adage goes, the only certainties in life are death and taxes. Even then, we cannot predict the time of our demise or foretell changes in the tax law. Therefore, we must venture into the world of uncertainty. The presence of uncertainty in public policy analysis means that decisions must be made *before* we know which one of many conceivable and perhaps even inconceivable outcomes will come to pass. While deciding on a possible outcome, we assign probability values, as we do when we assess the odds of winning the lottery, the chances of a rainstorm, or the behavior of the stock market. The sources of probability measures are either objective (based on actual facts) or subjective (based on personal judgment). Let us discuss these two sources of probability.

OBJECTIVE PROBABILITY

The **objective** measure defines probability as the ratio of the occurrence of a given event from a finite number of possible outcomes. Therefore, objective probability is also known as **relative frequency**. Say that 245 students with a certain range of grade point averages and SAT scores applied for admission to a public university and that 35 of them were selected. We can calculate the probability of a student's being admitted as

$$p = \frac{35}{245} = .143$$

In other words, the probability (p) of a student in that particular academic achievement category gaining admission is 14.3 percent. The formula for calculating objective probability by the measure of relative frequency can be written as

$$p = \frac{\text{Observed frequency}}{\text{Total frequency}} \quad (7.1a)$$

However, defining probability by observed frequency alone is problematic. If I ask you, for example, what the probability is of flipping heads in a coin toss, you will notice that it is impossible to use the above definition of probability, because we cannot observe all the coin tosses and record their outcomes. In that case, we may define objective probability as

$$P = \frac{\text{Target outcome}}{\text{Total number of possible outcomes}} \quad (7.1b)$$

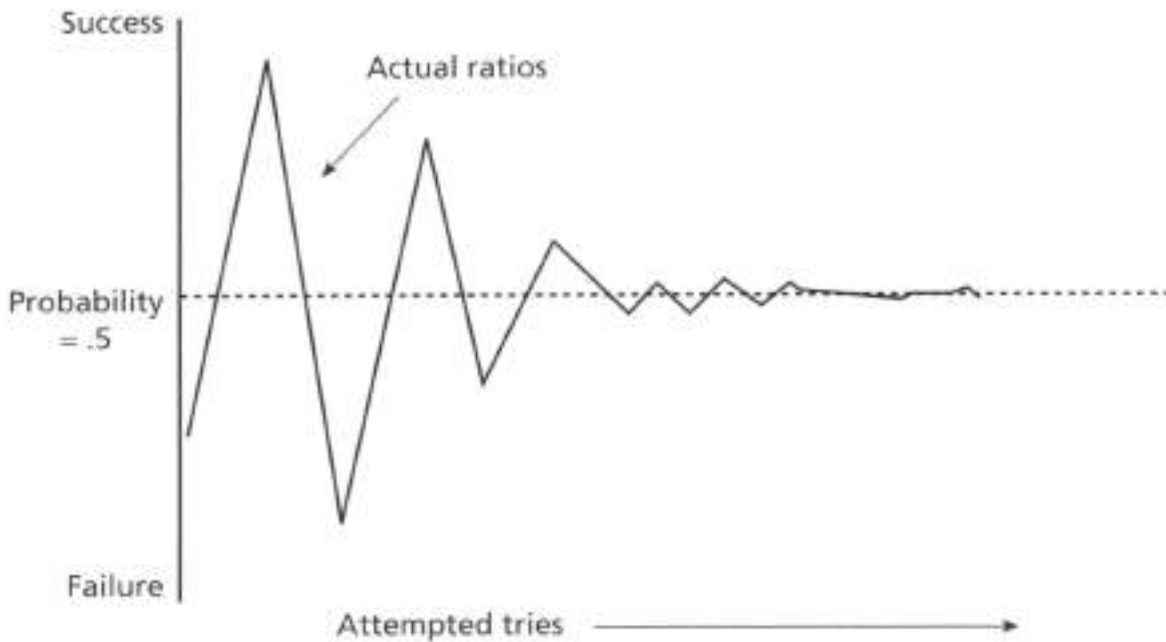
Suppose we toss a coin looking for one specific outcome, heads, out of two possible outcomes. Using formula 7.1b, we can calculate that the probability of flipping heads is $1/2 = .5$. Similarly, the probability of getting a 5 in the roll of a die is $1/6 = .17$.

We cannot keep on tossing the coin an infinite number of times to prove that the probability of flipping heads is $1/2$. Thus you must recognize that calculated probability in actual practice is only an approximation of the "true" probability. As the number of experiments becomes larger and larger, approximation poses less of a problem (see "A Case in Point"). Let us take a moment to explain this extremely important concept properly.

It is obvious that a weather forecast of a 50 percent chance of rain does not imply that it would rain half the day or that it would rain every half hour. In fact, it will either rain or it will not (even drizzle should count as rain). Interpreting such a forecast is the same as interpreting the outcome of a coin toss. Since by tossing a coin you will either win or lose, you should interpret the probability measure as if you were to toss the coin an unlimited number of times. You will see that as the number of tosses increases, the ratio of heads to the total number of attempts approaches .5 (see Figure 7.1). Suppose we are tossing a coin in groups of four tosses. After each set of tosses, we record the number and average it over the previous ones. Suppose that on the first try, we flip heads one time and tails three times. This gives us a ratio of 1:3. Suppose that on the next try, we flip heads twice and tails twice. Averaging the total number of tosses, the ratio of heads to tails is 3:4. If we continue to do this, we will see a curve similar to the one presented in Figure 7.1. In the beginning, the fluctuations may be wide, but as the number of tries increases, the average approaches the "true" probability figure of .5.

This is the basis of objective probability. Embedded in this simple example of a coin toss are some of the most important assumptions and axioms that form the basic building blocks of statistical reasoning. A proper understanding must begin with an examination of these properties of objective probability. First, the events in a probability experiment (in this case, each toss) must be **independent**. Perhaps the most common cause of misinterpreting probability is the question of independence. Suppose you have placed a bet on a coin toss. You are calling heads, and you have called it correctly four times. On the fifth try, should you call heads again, or should you try to be smart by calling tails? Because the tosses are independent, what happened last time or any time before has no bearing whatsoever on what may happen this time.

The second assumption of objective probability is that the events are **mutually exclusive**, meaning that they cannot take place simultaneously. For example, we cannot get both heads and tails on a single try. But it is entirely possible to have both rain and sunshine within a single day. Therefore, whereas the outcomes of a coin toss and the roll

FIGURE 7.1 Probability and the Number of Tries

of a die are mutually exclusive, rain and sunshine are not, unless we define the outcome as “having rain any time during the day.”

Finally, the probability measures are **asymptotic**. That is, the probability of flipping heads with an unbiased coin is .5. However, that does not imply that we will get exactly two heads and two tails in four tries. It simply means that if we repeat the experiment many times, as the number of tosses increases, the ratio will approach .5. Similarly, going back to our previous example, we must interpret the meteorological prediction as follows: if the weather conditions (temperature, barometric pressure, wind condition, humidity, and so forth) become the same as the condition we are having at this time, over a long period of time, this combination of conditions will produce rain half of the time.

Probability Distribution

When we assign probability to a series of independent and mutually exclusive events, we get a **probability distribution**. Suppose we want to find the probability of the size of a social worker’s caseload on a given day. We record the number of cases each day—say, for 100 days. Then, by dividing each day’s frequency by the total number of arrivals, we can construct the probability column of the arrival of the cases. These hypothetical data are presented in Table 7.1 and Figure 7.2. By looking at the table, you can tell that the probability that the social worker would handle just one case is 5 percent, two cases is 10 percent, and so on. The third column of the table shows the cumulative probability totals. Called a **cumulative probability distribution**, the numbers are the probability of having a value less than or equal to a specific value of x . The cumulative probability distribution data reveal that the probability of the social worker’s having four or fewer cases on a given day is 47 percent.



A CASE IN POINT

Probability in Baseball: Revenge of the Nerds

At first glance, it is hard to believe that Bill James (b. 1949) is responsible for winning so many baseball games; he is pudgy, middle aged, and in sum the antithesis of an athlete. His secret: a deep understanding of probability theory and the statistics that explain the outcome of the game. Every sport is ultimately about probability. Even if all the “best” players are on a team, they are not assured victory every time.

Among all the professional sports, baseball is perhaps the best suited for statistical testing, because of the huge number of chances players get during the season. The regular season of major league baseball consists of 162 games, each with at least nine innings. Assuming each player gets 3 shots at the plate per game, a starter can appear at the plate 486 times in a season. Therefore, keeping statistics has been a preoccupation of baseball fans ever since 1845, when Henry Chadwick invented the box score (a brief numerical summary of the outcome of a baseball game that lists each player's contributions—times at bat, hits, runs batted in, and so on—and describes how runs were scored over the course of the game).

Bill James, however, took this score keeping to a new level.¹ By looking at the detailed statistics, he would calculate the probability of success and test his hypotheses with the actual outcomes of ball games. Over the years, James dispelled many bits of conventional wisdom about the game, but perhaps his most important contribution was related to walks (in which a batter proceeds to first base after four pitches are thrown outside the strike zone, referred to as balls). Old-fashioned coaches and managers never kept count of the number of walks batters had; instead they were fixated on the number of hits. However, Bill James asked the most basic

Knowledge of the mean and standard deviation of a probability distribution provides us with a good deal of information about the nature of the distribution. Most important, this information gives us some of the most powerful analytical tools in the field of applied statistics. Probability distributions can take on an infinite number of shapes. However, for analytical purposes, a perfectly symmetric distribution, known as the **normal distribution** (or, in common parlance, the bell curve), serves as the fundamental building block. Again, a symmetric distribution can also take on various shapes. In Figure 7.3, I have drawn three normal distributions with different means.

Symmetric distributions can also have the same mean and yet be different. In Figure 7.4, the three distributions, despite having the same mean, have their relative levels of dispersion, measured in terms of standard deviation, σ . Clearly, of the three normal distributions, the tall and skinny one has the lowest standard deviation, whereas the short and fat one has the largest dispersion. Because of these variations, a **standard normal distribution** is defined as the one whose mean is equal to 0 and standard deviation is equal to 1. This standard normal distribution serves as the “ideal type,” or the benchmark against which the probabilities of an uncertain world are measured.

question: what is the purpose of batting? To him, the entire purpose of batting was to score runs. When he analyzed the records, he came up with the following formula:

$$\text{Probability of runs created} = \frac{\text{Total bases} \times (\text{Hits} + \text{Walks})}{\text{Plate appearances}}$$

When he plugged in the variables, he was spot on in explaining the total number of runs made. As you can see from his formula, a batter's ability to stay on the field and avoid getting out is just as important as getting hits.

James also dispelled many other old ideas about the game. For instance, he showed that sending in "clutch hitters" in pressure situations made no difference in the outcome of the game. James also paid special attention to the men on the field; he noted that although the number of catches can indicate a fielder's capabilities, a wider measure of the number of assists to get the other side out also had to be taken into account. James called this the "range factor."

On the basis of his findings, teams started keeping records that were more detailed than standard box scores. Since major league baseball is really a business, many other decisions—such as the position of the outfield wall, the percentage of left-handed versus right-handed batters, and the value of a player on the basis of his age—have been guided by James's calculations.

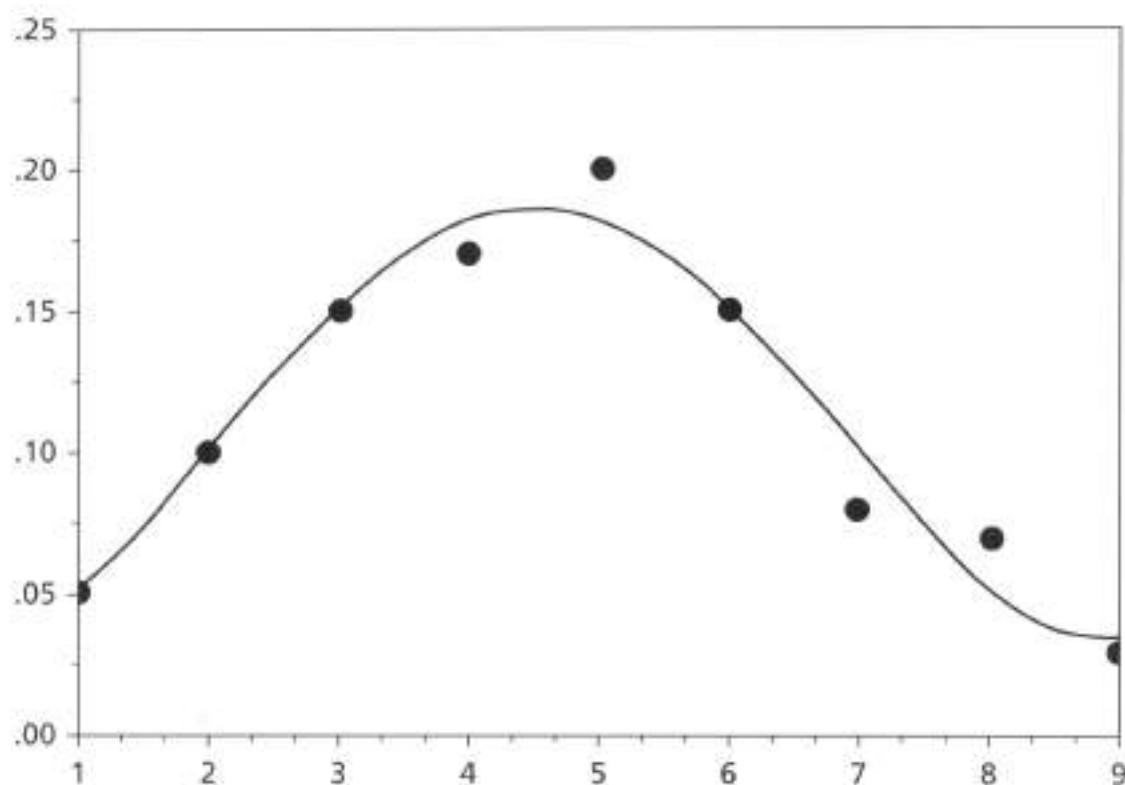
Discussion Points

1. What does Bill James's story tell us about the use of numbers in decision making?
2. Can you think of areas in the sphere of public policy in which gathering more information can help us achieve greater efficiency?

TABLE 7.1

Probability and Cumulative Probability Distribution of Caseloads for a Social Worker on a Given Day

Number of cases (x)	Observed frequency	Cumulative probability distribution	Probability distribution
1	5	.05	.05
2	10	.10	.15
3	15	.15	.30
4	17	.17	.47
5	20	.20	.67
6	15	.15	.82
7	8	.08	.90
8	7	.07	.97
9	3	.03	1.00
Total	100	—	—

FIGURE 7.2 Plot of Probability Distribution

The beauty of a standard normal distribution is that the distribution is standardized, as the name suggests. That is, we are able to measure any segment of it with the utmost precision. Because for a standard normal distribution the standard deviation is 1, any distance from the mean can therefore be measured in terms of the standard deviation. For example, from Figure 7.5, you can see that 68.26 percent of all the observations in the distribution will fall within $+1\sigma$ distance from the mean. Similarly, the area within the boundaries of $+2\sigma$ will capture 95.46 percent of the observations; 99.74 percent of the observations fall within $+3\sigma$, and nearly 100 percent are captured within the bounds of $+4\sigma$. The formula of the normal distribution enables statisticians to calculate any defined area under the curve.

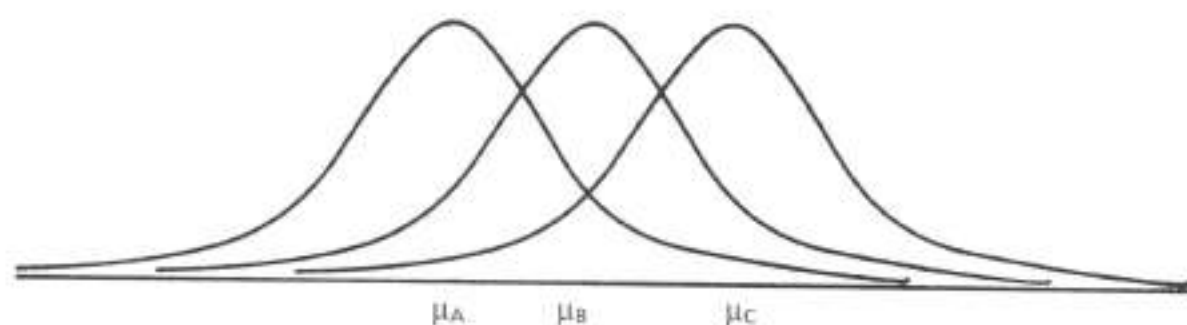
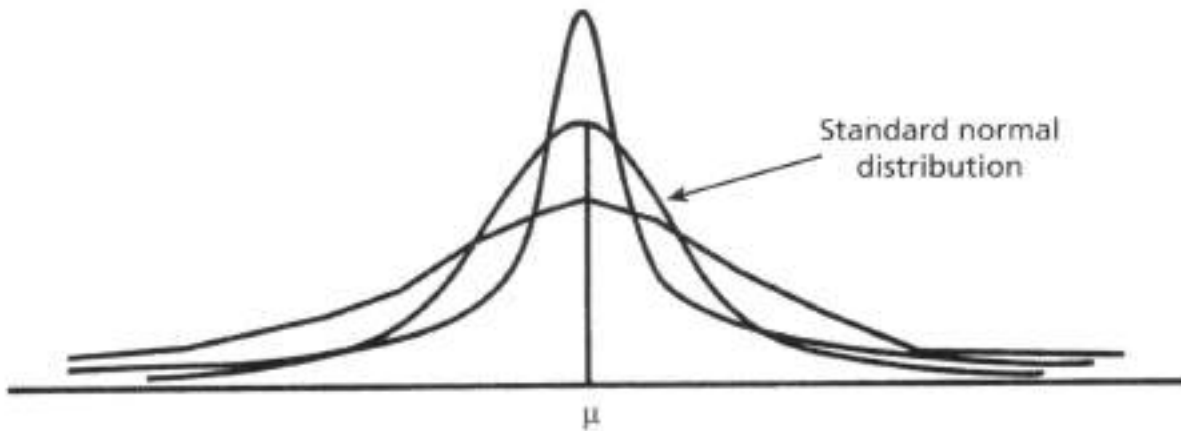
FIGURE 7.3 Normal Distributions with Different Means

FIGURE 7.4

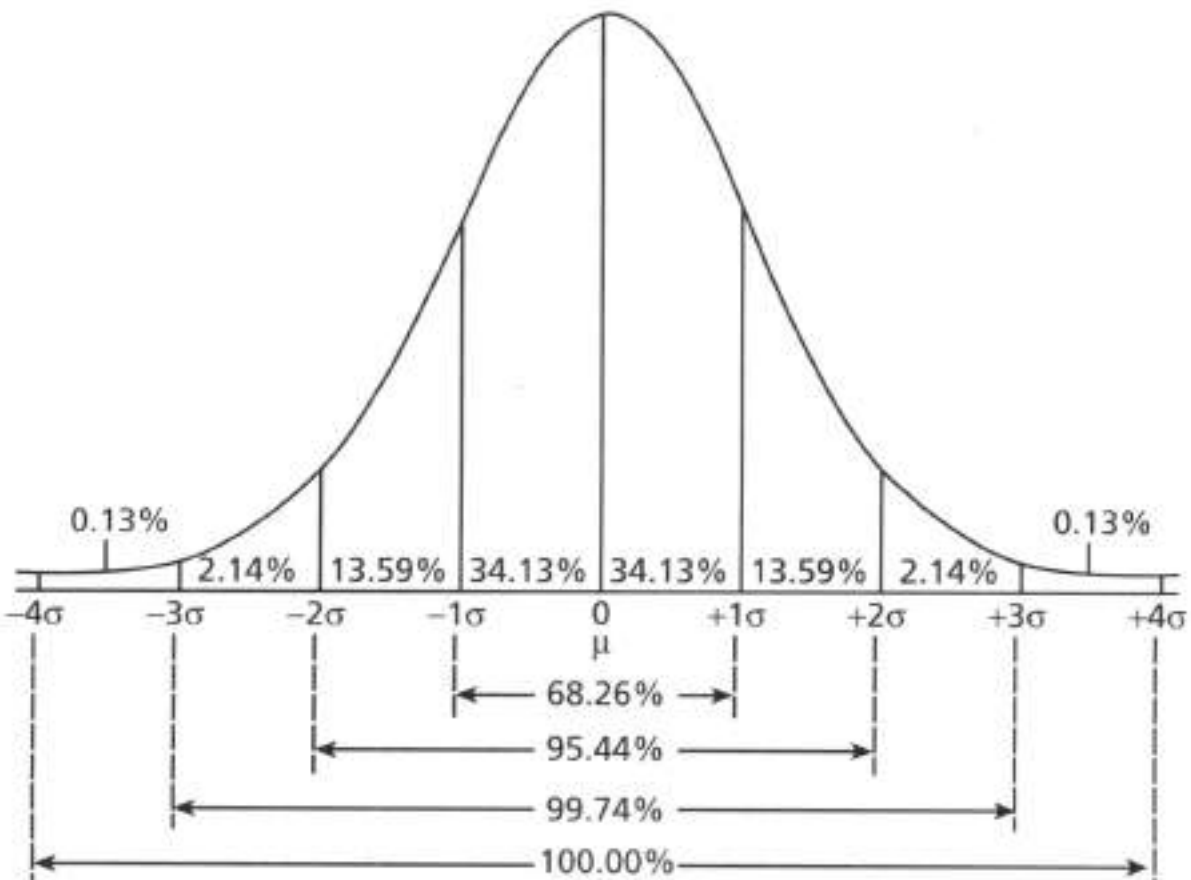
Normal Distributions with Various Levels of Dispersion



The standard normal distribution is, of course, a theoretical concept. The **central limit theorem**, perhaps the most remarkable theorem in the field of mathematics and statistics, bridges the gap between the theory and its practical use.² This theorem is also often referred to as the **law of large numbers**. An example will help you understand the compelling nature of this theorem. We know that human beings come in all shapes and sizes. Therefore, if we plot the percentages of males in the United States falling in various

FIGURE 7.5

Standard Normal Distribution



groups of heights, we will see a distribution with a strong central tendency and relatively few observations at the extreme ends of the spectrum (extremely tall and extremely short). This plot will not show a perfectly symmetric distribution, but we can get a standard normal distribution if we follow a simple procedure. Suppose we already know the mean (μ) and the standard deviation (σ) for the height of the U.S. male population. We collect ten males at random, note their average (mean) height, and let them go. We then collect another group of men, note their average height, and let them go. If we perform this operation many times, we will get a series of numbers showing sample means, which we may call $\bar{X}_i$. The reason for the subscript i is that we are referring to a series of means, not just one single value of a mean.

The central limit theorem tells us that if we subtract the sample means from the population mean and then divide by the standard deviation, as we repeat the experiment we will obtain a standard normal distribution, regardless of the shape of the original distribution. Suppose we know that the average height of an adult American male is 68 inches and the standard deviation is 2.0. Now we collect the averages of three groups of ten men and see that their average heights are 72, 73, and 65 inches. We then perform the following operations:

$$\frac{72 - 68}{2} = 2.0; \frac{73 - 68}{2} = 2.5; \frac{65 - 68}{2} = -1.5$$

As the number of samples increases, the series will have a mean of 0 and a standard deviation of 1. We can write this with the help of mathematical symbols as follows:

$$\frac{\bar{X}_i - \mu}{\sigma} \approx (0, 1) \text{ as } i \rightarrow \infty \quad (7.2)$$

We read expression 7.2 as follows: the series derived by subtracting the population mean from sample means, and divided by the population standard deviation, will be distributed with a mean of 0 and a standard deviation of 1.³ Another example will clarify the enormous practical implication of this property.

Several years ago, two neighboring countries in a bitter border dispute claimed the skeletal remains of some prehistoric settlers that had been found in the area. The discovery caused excitement because the racial identity of the ancient inhabitants could decide the current dispute. Thus, researchers took the average cranial measurement of the newly discovered skeletons. Thanks to previous anthropological studies, the average cranial measurement and the standard deviation of the present stock of occupiers of the land were known. It was decided to test the hypothesis that the people whose skeletons were discovered came from the same group of people as today's inhabitants.

However, in mathematics it is often impossible to prove the *existence* of a relationship. If we state that all A's are B, it would be impossible to check every A to prove the relationship. Instead, it is often simpler to nullify an opposite hypothesis in its negative form, such as "there are some A's that are not B." If we do not find any example of an A not being a B in our sample, then as the size of the sample increases we become more and more confident in our assertion that there is no A that is not B. In statistical terminology, this alternate hypothesis is called a **null hypothesis**. We call it "null" (defined

by one dictionary as “none,” “invalid,” or “void”) because while testing, the analyst is usually interested in the hypothesis being false.

Figure 7.5 can help explain the idea behind this test. If the cranial average of the skeletal remains is very close to that of the present group, we can assume that they both belong to the same racial stock. However, if the average falls outside the 2σ measurement from the mean, the probability distribution tells us that there is only about a 4 percent chance that these people belonged to the same racial group, because nearly 96 percent of all the observations (in this case, people of this particular race) will fall within this range. The process we just described is known as **hypothesis testing**, part of the fundamental construct of analytical reasoning.

Hypothesis Testing and Confidence Intervals

The theoretical framework derived by the central limit theorem can be used to test the probability of a hypothesis or to estimate a band within which we can reasonably expect to find the “true” value. Suppose the average drop-out rate in a school district is 15.5 percent, with a standard deviation of 1.05. Within the district, one particular school is being touted as exceptional in its achievement of a lower drop-out rate. Last year, this school showed only a 15.3 percent drop-out rate in a graduating class of 100. Is this truly exceptional or merely a result of chance factors? The basis of this question forms the core of hypothesis testing.

Hypothesis testing starts with a question. In the preceding example, the question is whether the drop-out rate of the school in question is significantly smaller than the average for the district. The second step in hypothesis testing is the formulation of a specific null hypothesis. Statistical convention generally expresses the research hypothesis as H_1 and the null hypothesis as H_0 .

These two hypotheses can then be written as follows:

H_1 : *The school's average is significantly lower than the district average.*

H_0 : *The school is not exceptional, and the difference between its mean and the average of the school district can be explained by chance.*

Since it is extremely important to construct a null hypothesis correctly, I offer another example of setting up a research question and a null hypothesis. What to do with juvenile delinquents deeply concerns society. These youngsters should be punished for their offenses, yet sending them to prison might transform them into hardened criminals. Suppose we are attempting to find out if setting up military-style boot camps for first-time juvenile offenders has any bearing on their future criminal activities. We have collected data by tracking two groups of recently paroled offenders: those who were sent to the boot camps and those who were sent through the usual criminal justice system. We can form the two hypotheses as follows:

H_1 : *Boot camps do reduce the probability of a juvenile's future conviction rate.*

H_0 : *An individual's enrollment in a boot camp has no bearing on his or her future conviction rate.*

In the third step, we must determine the **level of significance** for rejecting the null hypothesis. That is, by using the standard normal distribution, we must determine the

level of certainty we must seek for its rejection. This level can be set at 90 percent, 95 percent, or even 99 percent. The level of significance in statistical jargon is often denoted as α (alpha). Suppose, for testing our hypothesis, we have set the α level at 90 percent. If we want to include 90 percent of the observations in a symmetric distribution, the remaining 10 percent must be divided equally between the two tails of the distribution. Therefore, in this case, when we want to know if the sample mean (the average for the particular school) is significantly *less* than the population average (the district average), we set the α value at half of .10, or .05. This is called a **one-tailed test**, shown in Figure 7.6. The probability value corresponding to the predetermined level of α is also called the **critical value**, since it determines the threshold for rejecting the null hypothesis. If we set the critical value at the 95 percent level of probability, we want to be 95 percent certain that what we see in our sample result has only a 5 percent chance of being caused by chance alone. If the sample value falls below the critical value, we will not be able to reject the null hypothesis.

Z test. Let us go back to our example of the school district. We know the mean and standard deviation for the district, so it is a simple matter to calculate how far the individual school's record is from the district average, measured by units of the district's standard deviation. We can then compare this number with a theoretical distribution for rejecting or not rejecting a null hypothesis. If the number of samples is large and we know the population mean and standard deviation, we can derive the value for a **Z distribution**, given by the formula

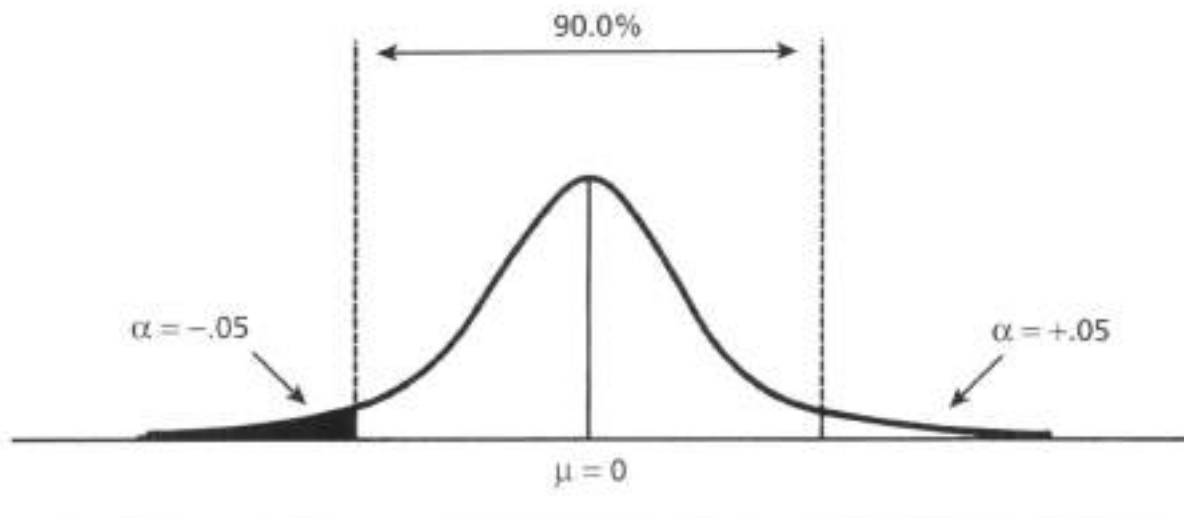
$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \quad (7.3)$$

where n is the number of observations in the sample, which in this case is 100 (the number of graduating seniors).

$$\begin{aligned} Z &= \frac{15.2 - 15.5}{1.05 / \sqrt{100}} \\ &= \frac{-0.3}{0.105} \\ &= -2.86 \end{aligned}$$

Our calculated **Z score** is -2.86 , which means that the sample mean (the mean of the school in question) is 2.86 times the standard deviation away from the population mean. The corresponding critical value of the Z distribution is given in Appendix C. The table in Appendix C gives the area under the right-hand side of the standard normal distribution, corresponding to the various calculated Z values. To determine the area of the normal distribution below our calculated Z value of 2.86, go down the first column and find 2.8. Then move to the right across the row to locate the number corresponding to column .06. The number corresponding to the row 2.8 and the column .06 is .4979, which is the area under the normal distribution below $Z = 2.86$.

Notice that this number relates only to the right-hand side of the distribution. Since it is a symmetric distribution, the other half contains 50 percent of the distribution. Therefore, by adding .5 to .4979, we get .9979, the cumulative value of the total area

FIGURE 7.6 One-tailed Test

below the area of the Z value of 2.86. This implies that we can be 99.79 percent certain that the difference between the average district drop-out rate and the experience of the model school is not due to chance; this school is justified in its claim to have significantly reduced the drop-out rate.

Our calculated Z value has a negative sign. Because the Z distribution is symmetric, this sign is of no consequence. It simply points to the direction of the difference from the mean. After calculating the Z score, you can find the corresponding probability value. If it is greater than the standard level of significance that you have set (usually .05), you can reject the null hypothesis. If it is less, you cannot reject it. You may also set the critical level of the Z score value without setting the probability standard. If you set your critical value at .05, you can go down the Z -score table to find the approximate point at which the area under the curve covers 95 percent of the distribution ($Z = 1.64$). If your calculated Z score is greater than 1.64, you can reject the null hypothesis.

Confidence Interval. The results of the Z tables can also be used to develop what is commonly known as the **confidence interval**. Clearly, the “true” values of a population distribution are unknowable. For instance, if we want to know the average height of American females, we have to measure every female in the United States, which is an impossible task. Therefore, we will have to infer the average height from sample results. Sample results converge with the actual value when the sample size becomes extremely large. Unless we want to continue taking larger and larger samples, it would be safer to express the population value with a band or an interval. Without saying that the average is 5 feet 6 inches, we will say that the average will lie between, say, 5 feet 3 inches and 5 feet 8 inches. If we express this band as a probability (for example, “We are 95 percent certain that the average height falls within this range”), then we call it a confidence interval.

A confidence interval tells us the range within which we can expect the population mean to fall with a certain level of confidence or probability. When radio or television newscasters tell us poll results, they almost always include corresponding confidence intervals. We are told that “67 percent of people support the proposed legislation, which

has an error margin of plus or minus 3 percentage points." In other words, we are being asked to hold with a high degree of confidence (usually at a 95 or 99 percent level) that the actual percentage of people in the general population who support this legislation falls between 70 and 64 percent. To derive a confidence interval, we need to use a **two-tailed test**, because we are not sure of the direction of deviation from the sample value. Therefore, consider the logic of a critical value for rejecting the null hypothesis on the basis of the Z score once again: reject the null hypothesis if

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > Z_{.05}$$

or

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < -Z_{.05}$$

These two expressions tell us to reject the null hypothesis if the Z scores are greater than the stipulated value. Diagrammatically this means that if the Z scores fall outside the shaded area in Figure 7.7, reject the null hypothesis.

The two algebraic forms can be rewritten as follows: reject the null hypothesis if

$$-Z_{.05} > \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > Z_{.05} \quad (7.4)$$

When we want to construct a confidence interval, we choose not to know a specific estimate for the population value μ . Instead, by manipulating equation 7.4, we can write

$$-Z_{.05} \times \frac{\sigma}{\sqrt{n}} > \bar{X} - \mu > +Z_{.05} \times \frac{\sigma}{\sqrt{n}} \quad (7.5)$$

which is again rewritten as

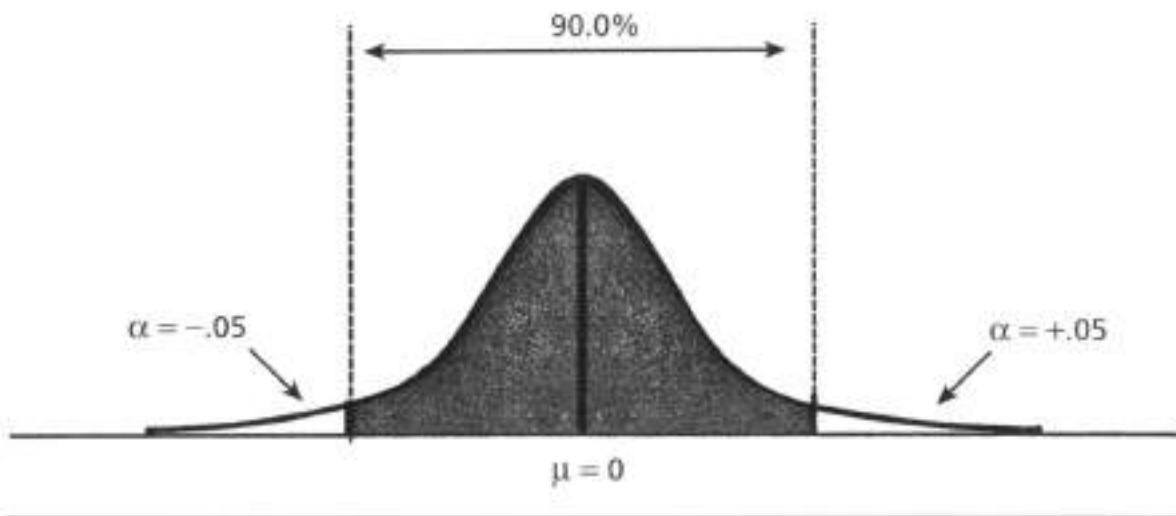
$$\mu - Z_{.05} \times \frac{\sigma}{\sqrt{n}} > \bar{X} > \mu + Z_{.05} \times \frac{\sigma}{\sqrt{n}} \quad (7.6)$$

Equation 7.6 gives us the confidence interval for our population (μ). In this equation, we set the value for the desired level of confidence, which in this case is 10 percent, or 5 percent on either side of the curve. By looking at the Z table, we can see that the value of $Z_{.05}$ is approximately 1.65. We know that the population standard deviation (σ) equals 1.05, n equals 100, and the population mean (μ) equals 15.5. Therefore, the confidence interval for our population mean is calculated by inserting these numbers in equation 7.6:

$$15.5 - 1.65 \times \frac{1.05}{\sqrt{100}} > \bar{X} > 15.5 + 1.66 \times \frac{1.05}{\sqrt{100}}$$

or

$$15.33 > \mu > 15.67$$

FIGURE 7.7 Two-tailed Test and Confidence Interval

In plain English, the results mean that we can expect 90 percent of all the schools in the district to exhibit drop-out patterns that will vary between 15.33 and 15.67 percent. This information can be of great help to administrators in evaluating the performance of individual schools in the district. Those that exceed a drop-out rate of 15.67 percent are doing significantly poorly. Those that have drop-out rates less than 15.33 percent are doing significantly better than the district at large.

t test. The *Z* test assumes that we know the population standard deviation and that the number of observations in the sample is large (usually more than 30). However, both of these assumptions are frequently violated in hypothesis testing. In such a case, the *Z* table becomes inappropriate. W. S. Gossett, writing under the pseudonym Student, devised another test for these kinds of problems, which has since been called the ***t* distribution**. The results of this distribution are presented in a *t* table (Appendix D). The formula for the use of *t* statistics is similar to that of *Z*, and with a large enough sample, the two will coincide. The formula for the *t* distribution is given as

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n - 1}} \quad (7.7)$$

Notice that this new statistic (*t*) differs from *Z* in that its denominator contains the sample standard deviation (*s*) instead of the population standard deviation (σ), and we take the square root of the number of observations in the sample ($n - 1$). The subtraction of 1 from *n* gives us the **degrees of freedom** (*df*). One can easily explain the concept of degrees of freedom, but its use in statistics is not exactly intuitive. The term *degrees of freedom* refers to the number of free choices one can make in estimating a number. For example, say that the sum of two numbers, *X* and *Y*, is equal to 10. There are, of course, infinite counts of two numbers that add up to 10. However, if I choose the value of one number (say, 27), there is only one number (−17) that will result in a sum of 10. Therefore, in estimating the sum of two numbers, we have the free choice of only one number ($2 - 1$). We call it the degree of freedom, in this case, 1. By extending this logic,

we can see that if twenty numbers add up to 10, we will have 19 degrees of freedom; after we have chosen nineteen numbers, the twentieth number is automatically determined. Since we are attempting to estimate the population value with samples, we use $n - 1$ degrees of freedom. Therefore, it is important to remember to subtract 1 from the number of observations for each set of samples. So to calculate the t value, which is done on one set of samples, we use $n - 1$ degrees of freedom. In the following section, we will compare two sets of samples. For that problem, we will subtract 1 for each of the two sets of samples, n_1 and n_2 . Thus, the degrees of freedom for the t distribution will be equal to $(n_1 + n_2) - 2$.⁴

Let us go back to the example of the discovered remains of inhabitants of the disputed land. Suppose we found the skulls of seventeen adults, with an average circumference of 21.5 inches and a sample standard deviation of 1.4. The average for the present-day population is 23 inches. To test the null hypothesis that these ancient people belonged to the same race, we set up the t score as

$$\begin{aligned} t &= \frac{21.5 - 23.0}{1.4/\sqrt{17-1}} \\ &= \frac{1.5}{1.4/\sqrt{16}} \\ &= 4.29 \end{aligned}$$

Suppose that we want to evaluate the null hypothesis at the .01 level (alternatively, the 99th percentile level), which as a one-tailed test is $\alpha = .005$. Now we can look up the t table (see Appendix D). By going across the columns, pick out the .005 level. Then come down along the rows until this level is matched with 16 degrees of freedom (17 observations minus 1). This will give you the critical t value of 2.921. Our calculated t value is larger than the critical value, so we can reject the null hypothesis and conclude that it is highly unlikely that they belong to the same race of people. *It is important to note that in the empirical sciences, regardless of how large the Z or the t value is, we never establish the absolute "truth"; instead, we maintain that we can reject the null hypothesis.*

Comparison of Two Different Series of Numbers. In the previous examples, we tested the hypothesis that *a particular sample was different from the population*. However, we may frequently need to test whether two samples are *significantly different from each other*. In the previous example, we discussed whether the remains of the ancient settlers belonged to the racial stock of the present population. In contrast, suppose we have found two sets of remains from two different time periods. We want to know whether these two groups of people came from the same stock. In such a case, we will test the null hypothesis that

$$\bar{X}_1 = \bar{X}_2$$

Since we are not making any assumption regarding the relative size of the difference between the two means (that is, if $\bar{X}_1 > \bar{X}_2$ or $\bar{X}_1 < \bar{X}_2$), we must use a two-tailed test. If the number of samples (n_1 and n_2) is more than 30 for each series, we can use a Z statistic; otherwise we need to use a t statistic. To test the null hypothesis that the two

means are equal, the difference between the two means is converted to a standard score by dividing it by the standard deviation. That is,

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{x_1 x_2}} \quad (7.8)$$

where $\bar{X}_1 - \bar{X}_2$ is the difference between the two means, and $\sigma_{x_1 x_2}$ is the standard deviation of the sampling distribution of the difference between two means.

If the t statistic is applicable, its degree of freedom is given by $(n_1 + n_2 - 2)$, where n_1 is the number of sample observations in series 1 and n_2 is the number of sample observations in series 2. Since we have the observations of the two series, we know their means. However, the question arises, What is the standard deviation of the distribution of the means? If there are reasons to believe that the standard deviations of the two sample series are in fact equal, the standard deviation of the combined series is calculated as

$$\sigma_{x_1 x_2} = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (7.9)$$

If, however, there is no reason to believe that the two standard deviations are the same, then we should use the following formula:

$$\sigma_{x_1 x_2} = \sqrt{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}} \quad (7.10)$$

Of course, the question arises, When should we assume that the standard deviations are the same, and when should we not? The answer to this question may depend on your subjective assessment. However, the rule of thumb is that if the two means are coming from a similar population group, you may assume that the standard deviations are the same. For example, if you are comparing the average number of children per household in two suburbs of the same city that have similar socioeconomic and demographic characteristics, you may safely assume that their standard deviations are the same. However, if you are comparing the mean of an affluent suburb with that of a poor center city area, you may have to assume that the standard deviations are different.

Let us consider a numerical example. Suppose we have data on drug abuse per 1,000 high school students in two school districts, A and B, with the characteristics shown in Table 7.2.

If we believe that the two standard deviations are equal, then using equation 7.10, we can estimate the pooled standard deviation as

$$\sigma_{x_1 x_2} = \sqrt{\frac{28(5.75)^2 + 16(6.05)^2}{28 + 16 - 2}} \sqrt{\frac{28 + 16}{28 \times 16}} = 0.589$$

By substituting this estimated number in equation 7.8, we get

$$t = \frac{25.0 - 20.6}{0.589} = \frac{4.4}{0.589} = -7.47$$

TABLE 7.2 Comparison of Means

	District A	District B
Mean	25.0 ($\bar{X}_1$)	20.6 ($\bar{X}_2$)
Standard deviation	5.75 (s_1)	6.05 (s_2)
Number of schools in the district	28 (n_1)	16 (n_2)

Since t has 42 degrees of freedom ($28 + 16 - 2$), the corresponding probability value for $\alpha = .025$ is approximately equal to 2.02. Because our calculated t score, -7.47 , falls outside the critical region of 2.02, we can reject the null hypothesis that the two means are equal. That is, drug abuse in district A is significantly lower than in district B.

If there is no reason to believe that the two standard deviations are the same, using equation 7.8, we estimate the pooled standard deviation to be

$$\sigma_{x_1x_2} = \sqrt{\frac{5.75^2}{(28 - 1)} + \frac{6.05^2}{(16 - 1)}} = 1.91$$

By plugging the value 1.91 into equation 7.8, we get $t = 4.4/1.91 = 2.30$. Again, this number is greater than 2.02, so we can reject the null hypothesis.

Although in this case both calculations gave us similar answers, it should not be assumed that the standard deviations in two series are equal.

Difference of Proportions Test. While doing a survey, you may face a situation in which you need to compare the proportions and percentages of the two sample populations. For instance, suppose you are conducting a survey to find out the attitude of citizens toward the police. Your survey indicates that 33 percent of the 320 whites polled agreed with the statement that “they do not have a great deal of trust” for the local police department, whereas 45 percent of the 267 minority respondents agreed with that assessment. To determine a policy of community relations for the police department, you want to know if the difference between the two communities is statistically significant. For this, you will have to use the following formula:

$$Z = \frac{\bar{p}_1 - \bar{p}_2}{\sqrt{\left(\frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2}\right)\left(1 - \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2}\right)\left(\frac{n_1 + n_2}{n_1 \times n_2}\right)}} \quad (7.11)$$

where

$\bar{p}_1$ = proportion of first subgroup,

$\bar{p}_2$ = proportion of second subgroup,

n_1 = sample population of first subgroup, and

n_2 = sample population of second subgroup.

For this example the values are as follows:

$$\bar{p}_1 = 0.38$$

$$\bar{p}_2 = 0.45$$

$$n_1 = 320 \text{ and}$$

$$n_2 = 267$$

When we substitute these numbers in equation 7.11, we get $Z = 1.71$.

The Z table tells us that the critical value of the one-tailed test is 1.645. Our Z value exceeds the critical value, and thus we can conclude that the minority groups have less trust in the police department. Hence, any public policy for community involvement must take into account this difference in perception.

The Chi-Square Test

The preceding examples show that if we are confronted with a problem such as determining whether a subgroup is significantly different from the larger population, we can use the Z test (provided, of course, that we know the population mean and the standard deviation). If we do not know the population values but want to know if two different groups are “truly” different from each other, we can use Student’s t test. In contrast to these two common forms of hypothesis testing, we frequently encounter a situation in which we may be interested in finding out whether an observed phenomenon is significantly distinct from its expected behavior.

Suppose the Metropolitan Transit Authority (MTA), which employs 300 bus drivers, has the task of training new recruits and periodically retraining older operators. The MTA used to operate an in-house driver training school; however, in a cost-cutting mood, it is experimenting with replacing trainers with a newly developed computer-based training method. By this method, trainees will sit in front of a computer terminal, go through each lesson, and at the end answer multiple-choice questions. If they pass, they will move on to the next section and will ultimately pass the entire theoretical portion of their training program. Management finds this computer-based program attractive: it offers cost savings on salaries and benefits, and the training school can remain open twenty-four hours a day, thereby allowing the drivers to come at their own convenience and train themselves. As a consultant, your task is to determine whether the system is acceptable to the drivers of the MTA. You conduct a survey that asks the drivers whether they are

1. enthusiastic about computer-assisted training
2. indifferent between the computer and the training school, or
3. upset about the change and would prefer the old training school.

Since fear of computers may be related to an individual’s gender, I decided to classify the trainees according to their sex and then look at the survey responses. I have presented the data in the form of a matrix, which is known as a **contingency table** (see Table 7.3). In this table, the numbers next to the word *observed* represent the actual numbers of responses corresponding to the attitudes toward computer-assisted training, classified by the gender of the respondents.

TABLE 7.3 Example of a Contingency Table

Sex of the trainees	Like computers	Indifferent	Do not like computers	Total
Female	Observed: 15 Expected: 25	Observed: 10 Expected: 13	Observed: 30 Expected: 17	55
Male	Observed: 120 Expected: 110	Observed: 60 Expected: 57	Observed: 65 Expected: 78	245
Total	Observed: 135	Observed: 70	Observed: 95	300

Below each observed frequency of responses is an entry for the expected response. This expected response is based on the hypothesis that there is no age bias in the acceptance of computer-assisted training. Thus, the two figures are exactly the same proportion as the two classifying variables (sex and attitude toward computers). This expected number is derived by multiplying the column total (known as the *marginal*) by the row total for each sex and dividing by the total number of trainees (300). The expected value for the female trainees who like computers is calculated as

$$\text{Expected value} = \frac{\text{Column marginal} \times \text{Row marginal}}{\text{Total frequency}} = \frac{135 \times 55}{300} = 24.75$$

The logic for this operation is that if preference for computers does not depend on the gender of the respondent, the women will prefer it in exactly the same proportion as all those who prefer the computers.

After we have determined the expected value for each cell, we can look into the difference between the actual and the expected values. Similar to the other probability distributions, the sum of the squared differences between actual and expected values is also distributed along a theoretical probability distribution, called χ^2 , or **chi squared**. This is written as

$$\chi^2 = \sum \frac{(f_o^i - f_e^i)^2}{f_e^i} \quad (7.12)$$

where f_o^i is the observed frequency of cell i , and f_e^i is its expected value.

Thus, we can write the formula as

$$\begin{aligned} \chi^2 &= \frac{(\text{Observed frequencies} - \text{Expected frequencies})^2}{\text{Expected frequencies}} \\ &= \frac{(15 - 24.75)^2}{24.75} + \frac{(10 - 12.83)^2}{12.83} + \frac{(30 - 17.41)^2}{17.41} \\ &\quad + \frac{(120 - 110.25)^2}{110.25} + \frac{(60 - 57.17)^2}{57.17} + \frac{(65 - 77.58)^2}{77.58} \\ &= 16.60 \end{aligned}$$

Like the t distribution, the χ^2 distribution varies with the degrees of freedom. For the χ^2 distribution, the degrees of freedom are calculated as (number of columns - 1) $\times$

(number of rows - 1). Thus, the degrees of freedom for our example are $df = (\text{columns} - 1) \times (\text{rows} - 1) = (3 - 1) \times (2 - 1) = 2 \times 1 = 2$.

Now we can look up the χ^2 distribution given in Appendix G. As you can see, for 2 degrees of freedom, the critical value at the .05 level is 5.99. Since our calculated χ^2 value, 16.60, is greater than this number, we can safely conclude that the results show a definite sex bias in the acceptance of computer-based training for the MTA bus drivers. You may notice that in this case, we did not need to know anything at all about the mean and the standard deviation of either the sample or the general population. This is the strength of this test and why it belongs to a class called **nonparametric tests**. Also, you should note that whereas t and Z distributions compare the difference between only two values, the χ^2 distribution measures the significance of all the cells jointly. Therefore, it is called a **joint probability distribution**.

Testing for Correlation: Pearson's r . In chapter 6 we discussed Pearson's r , or the correlation coefficient. As you know, the relation between two variables in the real world is seldom perfect. Therefore, in observing a less than perfect correlation, you may wonder if the calculated coefficient is statistically significant. One of the advantages of the correlation coefficient is that it can be used to test the hypothesis that two variables are correlated. To test this hypothesis, r must be converted into a t statistic using the following formula:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (7.13)$$

where n is the size of the sample.

For example, suppose the correlation between X and Y has been found to be .65, with a sample of 26 observations. The correlation is less than perfect, so you can formulate the null hypothesis that there is no correlation between the two variables.

By using the preceding formula, you can calculate t as

$$t = \frac{.65\sqrt{26-2}}{\sqrt{1-.4225}} = 4.18$$

Since there are 26 observations, the degrees of freedom for the t test are $26 - 1 = 25$. The critical value for t at a 5 percent level of confidence for 15 degrees of freedom is 2.060. Since our calculated t statistic is greater than this critical value, we can reject the null hypothesis.

The Risks in Hypothesis Testing. In testing hypotheses under uncertain conditions, we run the risk of (1) rejecting the null hypothesis when it is in fact true or (2) accepting a false hypothesis. A true hypothesis rejected as false incurs a **type I error**, whereas a false hypothesis accepted as true creates a **type II error**. The matrix shown in Table 7.4 best explains this dilemma.

Many public sector decisions affecting the health and welfare of citizens are based on the statistical reasoning of hypothesis testing. Therefore, the cost of an erroneous decision can be enormous. Suppose we are considering school funding on the basis of the

TABLE 7.4**Errors of Hypothesis Testing**

Decision	Hypothesis is true	Hypothesis is false
Reject hypothesis	Type I error	Correct decision
Accept hypothesis	Correct decision	Type II error

attendance records of pupils. A particular school shows a higher than average absentee rate, which puts it at risk for losing part of its revenue from the government. If we reject the hypothesis that the school is failing to provide a rewarding education system to keep students in class, and the hypothesis is indeed true, we will be wasting public funds on an inefficient (at least by this particular measure) institution. In contrast, if we accept the null hypothesis and find the school negligent of its duties, when in fact its attendance record is not worse than those of the other schools in the district, we will be inflicting undue pain on a school.

Therefore, we try to minimize these two errors in determining the decision rules. Unfortunately, these two errors are inversely related. That is, when we want to reduce type I error, we create more type II error. If we want to give the benefit of the doubt to the individual schools, we stand to encourage poor schools by failing to take punitive measures. In contrast, if we become extremely strict about attendance records in determining school funding, we inflict unnecessary punishment on some otherwise deserving schools.

In statistics, the type I error is the chosen level of significance, or the α level. Therefore, when we choose an α level of .05, we make sure that we will falsely reject a true hypothesis only 5 percent of the time. If we lower the significance level to .01, we reduce the chance of accepting a false hypothesis but increase the probability that we will inadvertently reject a true one.

The choice of a significance level depends on the type of problem under consideration and the consequences of rejecting a true hypothesis. The criterion for accepting a hypothesis is rather conservative, and given a choice, we would rather err on the side of the status quo than make a false move. An example may clarify the dilemma. In the early 1990s the U.S. Food and Drug Administration (FDA) came under fire from two different groups for two diametrically opposite reasons. On one hand, the FDA was accused of taking too long to test and approve for sale experimental drugs for AIDS. On the other hand, the FDA was heavily criticized for approving silicone gels for breast implants too quickly, without proper testing. If we are dealing with a problem that requires the utmost safety, we would be better off setting the level of significance as small as possible (say, at .001). However, it is a matter of common practice to set the confidence level at 95 percent ($\alpha = .05$) for most research and policy analysis.

Steps in Hypothesis Testing

1. Check theories relating to the comparability of the variables for testing.
2. State the null hypothesis clearly, making sure of the opposite hypothesis.
3. Specify the significance level of α .
4. Specify whether to use a Z or a t test.

5. Compute the value of the test statistic.
6. Draw your conclusion regarding the rejection or nonrejection of the null hypothesis.

SUBJECTIVE PROBABILITY

Ron and Chris are members of the parole board. They are participating in a parole board hearing of Rob. The board must decide whether Rob is a safe enough risk to grant him parole. Ron and Chris have different views. Ron sees Rob as a “high risk” for committing a serious crime, whereas Chris judges him to be a “low risk.” Both Ron and Chris are assessing the probability that this convicted felon will be in trouble with the law in the near future. However, in this case, they cannot use measures of objective probability, because each individual parole applicant poses a unique case. The decisions of Ron and Chris are prime examples of rendering judgment on the basis of **subjective** assessments of probability.

In our everyday lives, we all make subjective judgments about uncertain situations. We appraise the future by assigning chance factors or odds. Although in life it is often sufficient to state that “it is likely that the task will be completed within the next week” or that “there is a good possibility that the shipment will not arrive on time,” in professional work, these vague chance factors must often be expressed in numerical terms. It is more useful to know that an expert assigns a 30 percent chance of a certain legislation’s passing congressional scrutiny, or an 80 percent probability of a survey’s results being completed by next Thursday, than simply knowing the odds in qualitative terms (“a small chance” or “highly probable”). This assignment of numerical value gives more precision to the statement and as such can be compared with the assessment of some other expert.

Long neglected by serious researchers, the issue of subjective probability is increasingly becoming a topic of serious scientific inquiry. To Frank Lad, “The logic of probability merely formalizes the coherent implication of specified beliefs about events when beliefs are expressed in a graded scale from 0 (complete denial) to 1 (complete affirmation). . . . The ‘scientific method’ empowers the scientists with precisely the same inferential principles that it empowers the person on the street in the conduct of daily affairs, no more no less.”⁵ In fact, many studies in economics have found that subjective assessments of probability by experts can often rival the accuracy of those derived from sophisticated econometric models.⁶

Still, we may need to take a number of steps to ensure the accuracy of a subjective assessment. Returning to our example of the parole board, we may ask the following question: is there a measure of overall accuracy so that we can compare the subjective assessment of one board member with that of another, or the judgment of one parole board with that of another? Statisticians offer a number of different measures of accuracy, but let us discuss the most commonly used measure, the **quadratic score**.⁷ The quadratic score (Q) is given by the formula $Q = 1 - (p - A)^2$, where Q is the quadratic score of probability assessment, p is the assessed probability, and A is the actual outcome.

As you can see from this formula, if I assign 0 probability to an applicant and he stays out of trouble ($A = 0$), I have a perfect score of 1. However, if the parolee goes back to prison, my score becomes equal to 0.

TABLE 7.5**Risk Assigned to Parolees**

Parole board member	Parolee		
	<i>Rob</i>	<i>Joe</i>	<i>Phil</i>
Ron	High	High	Low
Chris	Medium	Low	Low

Let us explain the method with the example of our parole board members, Ron and Chris. They had categorized three successful parole applicants, Rob, Joe, and Phil, as high risk ($p = .8$), medium risk ($p = .5$), and low risk ($p = .2$). After a year, Rob and Phil had stayed out of prison ($A = 0$), whereas Joe had committed a serious crime ($A = 1$). On the basis of this information, we can construct the matrix shown in Table 7.5.

On the basis of our method, we can calculate the accuracy score for the two board members with regard to Rob as follows:

$$\text{Ron's score on Rob: } Q = 1 - (.8 - 0)^2 = .36$$

$$\text{Chris's score on Rob: } Q = 1 - (.5 - 1)^2 = .75$$

This measure magnifies the error in judgment by squaring the difference between actual and predicted values. By following this method, we can score the board members on each of their judgments and average the score. From Table 7.6, we can see that Chris has been more accurate than Ron. You may notice that although in this case, we can claim that Chris has been a better forecaster than Ron, you may not say for sure if that is purely by chance. It is of course possible to establish the statistical significance of the differences in performance; for that, however, we need many more observations. If we have a large enough sample size, we can use the techniques of hypothesis testing to find out the statistical significance.

Finally, we should mention the shortcomings of subjective probability. The attribution of probability scores by individual observers will always be open to question. We all assign probability values to uncertain events in our everyday lives. These attempted peeks into the future are based on our individual experiences and expertise, and thus personal biases of judgment influence our assessments of probability. To understand subjective probability, we must know the sources of biases that may obscure our judgments. I have discussed the psychological obstacles to critical thinking in chapter 5. They are equally applicable to the assessment of subjective probability.

TABLE 7.6**Quadratic Score Matrix for Parolees**

Parole board member	Parolee			Average score
	<i>Rob</i>	<i>Joe</i>	<i>Phil</i>	
Ron	.36	.36	.96	.56
Chris	.75	.96	.96	.89

KEY WORDS

- | | |
|--|---------------------------------------|
| Asymptotic probability (p. 151) | Null hypothesis (p. 156) |
| Central limit theorem (p. 155) | Objective probability (p. 149) |
| Chi-squared (p. 166) | One-tailed test (p. 158) |
| Confidence interval (p. 159) | Probability distribution (p. 151) |
| Contingency table (p. 165) | Quadratic score (p. 169) |
| Critical value (p. 158) | Relative frequency (p. 149) |
| Cumulative probability distribution (p. 151) | Standard normal distribution (p. 152) |
| Degrees of freedom (p. 161) | Subjective probability (p. 169) |
| Hypothesis testing (p. 157) | t distribution (p. 161) |
| Independent probability (p. 150) | t test (p. 161) |
| Joint probability distribution (p. 167) | Two-tailed test (p. 160) |
| Law of large numbers (p. 155) | Types I and II errors (p. 167) |
| Level of significance (p. 157) | Z distribution (p. 158) |
| Mutually exclusive probability (p. 150) | Z score (p. 158) |
| Nonparametric tests (p. 167) | Z test (p. 158) |
| Normal distribution (p. 152) | |

EXERCISES

1. What is an objective probability? What are the sources of bias in how we understand objective probability?
2. What is a subjective probability? What are the main sources of bias an individual faces in assessing subjective probability?
3. With an appropriate example, discuss the use of the central limit theorem as the foundation of statistical reasoning. Explain how it helps build a model of hypothesis testing.
4. The average SAT score for your state is 850, with a standard deviation of 65. The high school in your area has an average score of 980. Is this school an exception? Justify your answer.
5. From the information given in exercise 4, provide a confidence interval to develop a criterion for identifying schools with exceptionally good and unusually poor records.
6. Property tax is levied on the appraised market value of a property. However, it is often alleged that although the appraised values of lower priced houses are quite close to their market values (revealed when the properties are actually sold), the appraised values of higher priced houses are significantly lower than their market values. To investigate this allegation, the city of Masters conducted a study that found that for the forty lower priced houses, the ratio of assessed value to market value (A/M) was 0.89, with a standard deviation of 0.09. For a sample of forty higher priced homes, this ratio was 0.65, with a standard deviation of 0.15. Do you agree that the lower income homeowners in the city are carrying an unfair property tax burden?
7. The financial manager for the city of Masters is considering a switch from the existing money market fund to a new one. On the basis of the presented data series, do

you believe that the difference between the two is significant enough to warrant a change?

Year	Existing	New
1	7.5	6.7
2	8.0	9.3
3	7.9	5.5
4	8.5	10.9
5	7.0	5.3
6	9.2	12.6
7	8.2	6.7
8	7.4	4.3
9	8.8	10.7
10	8.0	12.9

8. The city of Masters has been conducting a survey of police responses to emergency calls. A similar city has been chosen for comparison purposes. The following table shows the distribution of response times during a typical week. Is there any reason to believe that there is any significant difference in the records of the two cities?

Distribution of Police Response Time

Response time (minutes)	Masters	Other city
1-3	23	45
3-5	62	72
5-8	41	59
8-10	38	43
10-15	16	32
15-25	3	21
25-35	0	13
35-60	0	7

9. Let us continue with the example of the MTA's new computer-assisted driver training program. Having examined the possible sex bias in the acceptance of the program, suppose you are trying to discover any possible age bias. Your survey results in the following table. Using a χ^2 distribution, determine whether age imposes an additional barrier to the use of computers in the MTA.

Age of the trainee	Like computers	Indifferent	Do not like computers	Total
18-25	30	25	10	65
26-45	65	30	60	155
46 and older	15	10	55	80
Total	110	65	125	300

10. Suppose we have observed ten individuals to establish the hypothesis that education has a strong correlation with income. Consider the following data, and test the hypothesis:

Education (years of schooling)	12	18	6	17	19	12	16	16	20	8
Income ($\times \$1,000$)	25	35	22	56	85	20	45	48	65	20

11. The crime rate in your state is 4,455 per 100,000 population, with a standard deviation of 289. The city in which you live has a crime rate of 3,990 per 100,000 population. Your police chief is claiming credit for a low crime rate. Is your city an exception? Assume that the city and the state have the same standard deviation.
12. Your state is considering a program for reducing recidivism among its convicted felons. Another state instituted a similar program ten years ago. The following table provides us with information on the two states. On the basis of these data, what advice can you give to the policymakers?

Recidivism per 1,000 Convictions

Year	Your state	State with a running rehabilitation program
1	155	134
2	160	183
3	153	110
4	172	193
5	140	145
6	188	252
7	166	132
8	151	86
9	214	116
10	260	98

13. Your city has been conducting a survey of incidents of lead poisoning among children in the city's school districts. For the purpose of comparison, you have chosen your neighboring city. The following table shows the distribution of toxic incidents among various age groups of school-age children during a six-month period. Is there any reason to believe that there is any significant difference in the records of the two cities?

Distribution of Lead Poisoning Incidents

Age of children	Your city	Neighboring city
5-6	54	53
7-8	51	43
9-10	44	38
11-12	24	29
13-14	10	9
15-16	5	1
17-18	0	1

14. The public health officials in your city are considering using a new type of flu vaccine that is being touted as a better preventive measure. However, the new vaccine costs three times as much per inoculation. On the basis of the following information, would you recommend the use of the new vaccine?

Comparison of Inoculation Using Two Types of Vaccines

Year	Number of flu incidents in your city per 1,000 inoculated (with the old vaccine)	Number of flu incidents in a neighboring city per 1,000 inoculated (with the new vaccine)
1	150	142
2	155	164
3	142	110
4	157	112
5	120	128
6	174	210
7	157	142
8	133	99
9	135	143
10	156	155

15. Write an essay on subjective probability. Suppose you are attempting to determine the odds that your favorite football team will win the Super Bowl. What kinds of biases might cloud your judgment?
16. In 2001 a school district was experimenting with a new policy to reduce absenteeism in its schools. The following table presents the data from the ten schools for two years. The 2000 data do not include the new policy. Using the methods of hypothesis testing, determine whether the district has been successful. Justify your methodology.

Comparison of Absenteeism

School	Rate of absenteeism in 2000	Regional average in 2001
1	4	9
2	3	5
3	4	6
4	5	4
5	3	9
6	9	9
7	2	7
	3	3

17. You are in charge of assessing the effectiveness of four parole board members. They assigned levels of high (H), medium (M), and low (L) risk for recidivism among parolees with corresponding subjective probabilities of .8, .5, and .2. You want to determine if their judgments were borne out by evidence after the parolees had been out of prison for one year. The number in parentheses, 1 or 0, after the name of each parolee indicates whether the parolee was back in prison within a year: the number 1 indicates that the individual went back to prison, and 0 indicates that he did not.

Parolee	Parole board member			
	Chris	Ron	Meg	Bob
A (0)	H	L	L	M
B (1)	H	M	L	M
C (0)	M	L	L	L
D (0)	H	L	L	M
E (1)	H	M	M	H
F (0)	H	H	L	M
G (1)	H	L	M	M
H (0)	M	L	L	L
I (0)	L	L	L	L
J (1)	H	H	L	L
K (0)	M	L	L	M
L (0)	H	M	L	L
M (0)	M	M	M	L
N (0)	H	H	L	L
O (1)	H	M	M	L
P (1)	H	L	L	M
Q (0)	M	M	L	L
R (0)	H	L	L	M
S (1)	H	M	M	M
T (0)	L	L	L	M
U (0)	M	H	L	H
V (1)	H	L	M	M
W (0)	M	M	L	L
X (1)	H	M	M	L
Y (0)	H	M	L	L
Z (0)	M	L	L	M

Notes

1. There are now many books in the market on Bill James and how he revolutionized baseball. Among them, I would recommend Michael Lewis, *Moneyball: The Art of Winning an Unfair Game* (New York: W.W. Norton, 2003); Scott Gray, *The Mind of Bill James: How a Complete Outsider Changed Baseball* (New York: Doubleday, 2006); and Gregory F. Augustine Pierce, ed., *How Bill James Changed Our View of Baseball* (Skokie, Ill.: ACTA Sports, 2007).
2. For an excellent discussion of the central limit theorem and its implication in the field of statistics, see John L. Casti, *Five Golden Rules: Great Theories of 20th-century Mathematics—and Why They Matter* (New York: John Wiley, 1996).
3. The fact that the mean of a normal distribution will be 0 can be easily shown. Since the sample average $\bar{X}$ becomes equal to its true population value μ , with infinite sampling the sum of $(\bar{X} - \mu)$ becomes equal to 0.
4. For the mathematical explanation of degrees of freedom, see Robert S. Pindyck and Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts*, 2nd ed. (New York: McGraw-Hill, 1981), 42–43.
5. Frank Lad, *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction* (New York: John Wiley, 1996), 13. Lad offers a comprehensive but demanding explanation of subjective probability.
6. See Harinder Singh, "Relative Evaluation of Subjective and Objective Measures of Expectations Formation," *Quarterly Review of Economics and Business* 30 (1990): 64–74; and Roger Frantz, *Two Minds: Intuition and Analysis in the History of Economic Thought* (New York: Springer, 2005).
7. J. Frank Yates, "Subjective Probability Accuracy Analysis," in *Subjective Probability*, ed. George Wright and Peter Ayton (New York: John Wiley, 1994), 381–410.

8

Sources of Data

The first question facing an empirical researcher is, Where do I get the necessary information to prove my hypothesis or to answer my question? Just as the first step of objective analysis is to define a project's goals or objectives, before you look for data, you must define precisely the problem you want to investigate. If your task is to get information on the effectiveness of a particular public policy, you need to plan precisely how you are going to go about your task. Then you must collect the information, which may or may not be readily available. Information available in published form is called **secondary data** (that is, data collected by someone other than the researcher). Data that you must collect on your own are **primary data**. In this chapter we will cover issues of measurement and then discuss primary and secondary sources of data.

WHAT ARE WE MEASURING?

We all like to quantify. We express the amount of rainfall, the speed of automobiles, even intelligence, in numbers. We collect statistical information on income, unemployment, crime, political violence, the extent of democratic values, and the level of authoritarian personality. As analysts, we aim to assign numbers to properties or attributes according to specific rules or measurements. However, controversies arise with regard to the accuracy of those measurements. Therefore, let us first discuss the various types of measurements and then concentrate on their relative merits as measuring units.

Types of Measurement

Phenomena to be measured can be assigned values according to nominal, ordinal, interval, or ratio scales. Of these measurements, the weakest form of assigning numerical value is the **nominal scale**. For example, while filling out the census form or application forms,

we are often asked to identify our ethnic background. Each ethnic group is assigned a number (1 through 5, say), a Roman numeral (I through V), or an alphabetic classification (A through E). Because these numbers and letters are not amenable to mathematical treatment, they do not give us any insight into their relative comparability, such as whether group B is closer to group A in any attribute or physical quality than is group E. For nominal scales, assigning numbers is devoid of any intrinsic meaning.

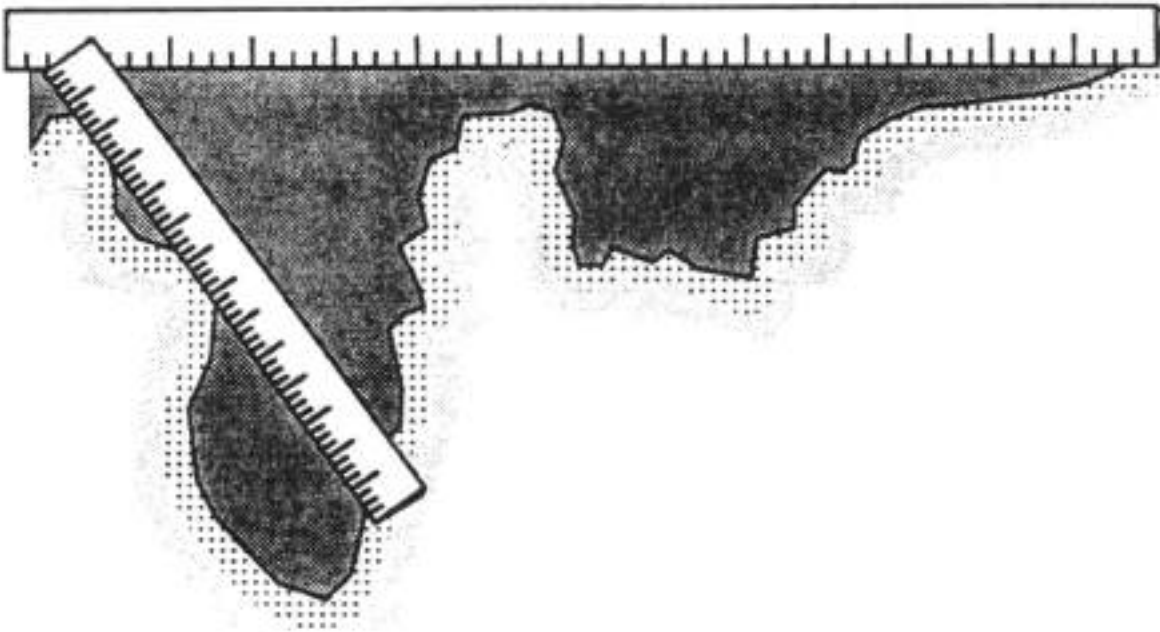
Some phenomena are comparable, so we can arrange them according to some quality but cannot tell precisely the distances among them. We may rank presidential candidates according to their political philosophies, but we may not be able to tell that candidate A is 2.35 times more conservative than candidate B. This process of ranking is called an **ordinal scale**. It is important to note that because numbers for ordinal rankings do not carry any specific meaning, the direction of their assigned values does not make a difference. Thus, among five candidates, the most conservative candidate may be assigned the value 1 or 5 without any consequence, as long as the ordering is in sequence. As you can see, ordinal rankings are stronger than nominal ones, since we can compare the relative position of the case in question (the "conservatism" of the presidential candidates in this example).

Arrangement using an **interval scale** offers us a greater amount of flexibility in comparing cases, both according to the arrangement of rankings and the actual distance between any two cases. If the candidates are assigned numbers on an interval scale, we can compare them by saying not only that A is more conservative than B but also that A is 2.35 times more conservative than B. Because of this desirable quality, the temptation to express orderings according to interval scales is strong. As a consequence, a great deal of effort has been directed toward constructing interval scale measures for various kinds of social, political, and economic phenomena. Various conservative and liberal congressional watchdogs assign numbers to individual members of Congress according to their voting patterns. In other studies, countries have been given numerical values for their development potential, degrees of democratization, and levels of political violence.¹

An interval scale, however, may not have a fixed and well-defined zero point, at which the quality we are measuring does not exist. In the progressive assignment of points (the higher the number, the higher the level of liberalism) to members of Congress, a senator with a score of zero must show no traces of liberal values whatsoever, an attribute rarely seen in human beings. In contrast, we may order various school districts by the percentage of minority enrollment, or airlines according to the ratio of delayed to total flight arrivals, and in each of these cases, the value of zero will have a specific meaning (a school district with no minority students or an airline that is always on time). This is called a **ratio scale**.

Accuracy of the Measuring Scales

When we try to quantify something, we need to know the accuracy of our measurements. If we want to measure a phenomenon that has a physical manifestation and only one dimension, our efforts are largely free of controversy. Measuring an individual's height or weight is done without any trouble, as long as we are measuring either height or weight. However, we are likely to run into a bit more trouble if we attempt to quantify the "largeness" of a person, a measure that must encompass both height and weight. Therefore, if we want to express a phenomenon that has no obvious physical manifestation, such as an

FIGURE 8.1 Measuring the English Shoreline

individual's intelligence, attitude, or quality of life, or the regional inflation rate, national growth potential, or international political instability, we are treading in dangerous waters. Even measuring strictly physical, single-attribute phenomena can sometimes be controversial. James Gleick gives an interesting example in a popular book on the mathematical theory of chaos.² In encyclopedias and other reference books, one frequently encounters various measures of national geography, such as the total square miles of land area or the length of shoreline. Gleick points out that the measurement of a jagged shoreline must vary with the size of the yardstick. When cartographers measure shorelines, they use approximations. Consider Figure 8.1. When I use a large yardstick, I miss all the little edges. However, by using a smaller stick, I can try to be more precise. In so doing, the measurement of the shoreline will register an increase, as this more precise instrument will include areas that were not measured before. Unfortunately, this smaller stick would not produce a perfect measurement, as it would miss areas smaller than itself. Clearly, as we keep reducing the size of the yardstick, we also increase the measurement of the total shoreline. In the end, when the yardstick becomes infinitesimally small, the measurement of the shoreline would be recorded as infinitely long.

We live in an imperfect world. Therefore, along with everything else, we must live with the shortcomings of our indicators. The only recourse we have is to be aware of their inherent biases and problems and to interpret them with extreme care. With these words of caution, we may examine the types of data and their sources.

PRIMARY DATA: CONDUCTING A SURVEY

Often, the kind of information you need is not available in archival records. In that case, you may need to collect information on your own. You can collect primary information through direct observation or through surveys. Traffic engineers routinely observe traffic

patterns for more efficient and safer road design. They can send actual observers or use counters that automatically record the number of vehicles and their speed. Public health officials develop regional statistics by compiling data from local hospitals, and water management departments collect information through laboratory testing of water samples from local rivers, streams, and other sources of drinking water.

Public policy researchers, however, most frequently use surveys to collect primary data. For example, local governments often rely on opinion polls for needs assessments, prioritization of issues, or to shape service delivery systems. If county authorities want to improve health care delivery to a community's elderly residents, they can commission a survey that assesses residents' needs through a series of specific or open-ended questions. A municipal service agency such as a rapid transit authority can design a survey for developing a more efficient mass transit system. A municipal government may also be interested in learning its constituents' views on a specific issue, such as a proposed ban on smoking in public places or the location of a proposed city library. The challenge for these and other surveys is how to collect unbiased information in a systematic way. Because it is often impossible to ask every member of a large group, we have to depend on surveys based on sampling theory. Sampling theory helps us determine the most cost-efficient way to derive unbiased information; more simply, it helps us design a survey.

Designing a Survey

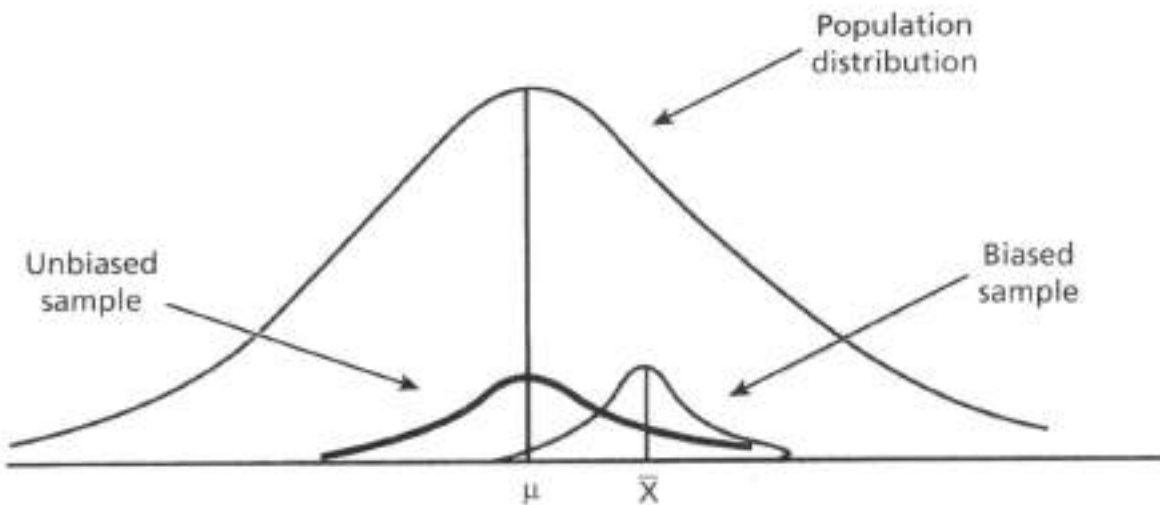
The starting point of any **survey design** is the question of **population**. The term *population* is defined as the entire target group from which the necessary information is to be extracted. Let us suppose that we want to determine whether the residents of a small town are willing to accept a proposed prison facility located on the outskirts of town. In such a case, we may define the population as all the city's adult residents. If we are trying to figure out the effects of a price hike on city buses, then bus riders will form the survey's population.

In a perfect world, with no constraints on time and money, it is desirable to have a complete enumeration, or a survey of each member of the population. However, in reality, we must confine our inquiry to a small group of representatives of the population, called the **sample population**. The single most important quality of the sample population is its representativeness. That is, how closely does this sample group represent the entire population? If it does not represent the population well, the sample population will give us biased (and therefore erroneous) information, as shown in Figure 8.2.

These **sampling biases** are characterized as either systematic or random. Through proper understanding of sampling theory, one can minimize systematic sampling errors, which result from inappropriate survey design. Random variations in the data are what cause random sampling errors, however. Such variations cannot be completely eliminated but can be managed through proper survey design. Let us discuss these in detail.

Systematic Errors in Sampling

Systematic errors result from biased sampling of a population. These biases have many sources. Some of the most common sources of systematic sampling biases are identified below.

FIGURE 8.2 Errors in Sampling

Nonrepresentative Sample. A nonrepresentative sample group produces biased inferences about a population. During the 1936 presidential campaign, the now defunct *Literary Digest* conducted a telephone survey. On the basis of its sample of New York residents, the magazine predicted a comfortable margin of victory for Alf Landon against Franklin D. Roosevelt. However, when the actual poll results came out, FDR had won a decisive victory. Why was the survey wrong? The prediction was incorrect because it was based on a sample that did not properly represent the population. Telephones were still a novelty in 1936, restricted to the relatively affluent. The telephone was not the tool of universal communication that it is today. Therefore, a sample of telephone owners was an inappropriate representative for the rest of the population.

The designer of a survey should always guard against this kind of systematic bias. For example, if you pick names of individuals from the property tax roll for a survey to determine the mass transit needs of a low-income neighborhood, your sample may diverge from your intended group—the neighborhood residents. Because most residents rent, rather than own, their homes in a low-income neighborhood, the exclusion of the renter population will cause the poll results to deviate significantly from those of the target population. To avoid this error one must have an intimate knowledge of the population and the existence of biased “cells” or subgroups within it.

Self-Selection Bias Within the Sample. Unfortunately, even correctly specifying the sample group will not always eliminate the problem of getting unbiased information. For example, suppose you carefully identified your sample group and administered a detailed questionnaire. When the results come in, you find out that they are still biased. This may happen because of self-selection within the sample group, when those who choose to respond to a survey turn out to be different from the rest of the population. Thus, going back to my example of mass transit needs in a low-income neighborhood, I may find that the respondents are disproportionately drawn from the English-speaking, better educated, older, and nonminority population. In such a case, their opinions will obviously not reflect the entire neighborhood. This self-selection bias may not be

restricted to individual respondents. The International City/County Management Association (ICMA) collects data on municipalities. However, its data sets are based on information voluntarily submitted by the municipalities. Such self-reporting can create a bias, because those local governments that are less affluent or are experiencing hard economic times may not have the personnel or the inclination to fill out a detailed questionnaire. Thus, the data sets presented in the ICMA yearbooks carry an appropriate warning.

When People Lie. If you ask a sensitive question, people often do not tell the truth. People may lie about their beliefs regarding race relations, gender equality, homosexuality, or other matters of private attitude because they are afraid to be politically incorrect or are mindful of social sanctions for unpopular opinions. This is a particular problem for AIDS researchers. To make a proper assessment of a community's needs, they must ask extremely sensitive questions regarding people's sexual habits. Unless they do so with the utmost care, their survey may produce inaccurate and biased results.

People are also typically reluctant to reveal their deeply held racial prejudices. As a result, time and again, forecasting the results of certain elections turns out to be incorrect. When Barack Obama became the first African American candidate for president in 2008, many pundits warned of the "Bradley effect," which posits that white voters hide their reluctance to support a black candidate from pollsters. Some saw Obama's defeat in the New Hampshire primary, in which he held a 10-point advantage in the polls leading up to the contest, as evidence of the Bradley effect, but his victory in the general election dissolved much of the buzz around this theory.

Collection Bias. Natural obstacles to reporting may create biases. In several border states the issue of illegal immigration has been rather controversial. Because immigration has a tremendous impact on a region's economy, the U.S. Census Bureau would like a more accurate enumeration of the illegal immigrants in the area and more detailed knowledge of their socioeconomic characteristics. However, for understandable reasons, illegal immigrants have been underreported.

Biases also result from the mode of collection. The *World Handbook of Political and Social Indicators* (1983), by Charles Taylor and Michael Jodice, makes a significant contribution to scholarly research by compiling international social and political data. Yet, as the authors readily admit, much of the political data are collected from the *New York Times* and a few other published reports. These newspapers are more likely to report on events that take place in Western nations than in some obscure country in Asia or Africa, so the data will contain systematic biases.

Errors Due to Observation. In the early 1920s, the physicist Werner Heisenberg performed a series of famous experiments in an attempt to observe the positions of subatomic particles. To ascertain the positions of these particles, they had to be charged with photon particles. But the very act of charging them caused them to change their original positions in a random fashion, thereby making it impossible to know where they were before. These experiments gave birth to what is known as the Heisenberg Uncertainty Principle, which simply states that for certain cases, no experimental design is available that would not create errors caused by the process of observation.



A CASE IN POINT

Errors of Observation: The Hawthorne Experiment

The fact that observation often alters the behavior of those being observed can significantly reduce the validity of research findings. The most celebrated example of this is the so-called Hawthorne experiment. In 1932 a group of researchers reported the results of an experiment they had been conducting for the past five years at the Hawthorne Works, a Western Electric Company manufacturing plant near Chicago. They discovered, by accident, that people generally seem to work better and are more productive when their performance on the job receives the flattering attention of a group of university researchers.

The researchers, headed by Elton Mayo and F. J. Roelithsberger, were originally trying to determine at what level the garment factory's illumination needed to be set to maximize production.³ They started with a very high degree of light and slowly reduced its intensity to see if it would influence productivity. The workers became aware of the fact that they were being observed. As the light got dimmer and dimmer, they took it as a matter of group pride to work even harder. Needless to say, their sensitivity to being observed spoiled the original intent of the research. However, the results of this experiment made scholars of organizational behavior aware of the strength of group identity as a motivating force in the workplace and started the so-called humanistic school of organizational behavior.

Discussion Points

1. Why should we be extra careful in designing our research, particularly when it comes to observing human behavior?
2. How can you maintain objectivity in observing human behavior or derive truthful answers through survey questions?

In an analogous way, social scientists often alter the behavior of the people they are observing. In other words, if we are told that we are being observed, our behavior may differ from that in our day-to-day lives. In a famous experiment conducted in the early 1970s, the psychologist Stanley Milgram demonstrated the power of authority figures in influencing individual behavior.⁴ His subjects were told that they were taking part in a test to determine the effect of punishment on memory. They were asked to give electrical jolts to another "subject," a man in another room, whenever he made a mistake. In fact, the voice from the other room was that of an actor and was taped for the experiment. They were also told to increase the voltage each additional time the "subject" made a mistake. To the amazement of the experimenters, the subjects continued to increase the voltage despite desperate pleas from the unseen voice, simply because they were told to do so by the laboratory psychologist. Nearly 60 percent of the subjects increased the voltage to the fatal level. Although the findings of this sensational experiment paint an unflattering picture of human nature, it is extremely difficult to say whether such blind obedience could be replicated on all of us in an open society.⁵

One of the ways to avoid the errors resulting from observation is to collect data on what is known as **revealed preference**. That is, although people's actual preferences are not apparent to an outside observer, they are revealed through their past choices. If you ask people if they will go to the polls in an upcoming election, those who may not vote may attempt to hide their actual intent because of a reluctance to admit their failure to carry out their civic duty.

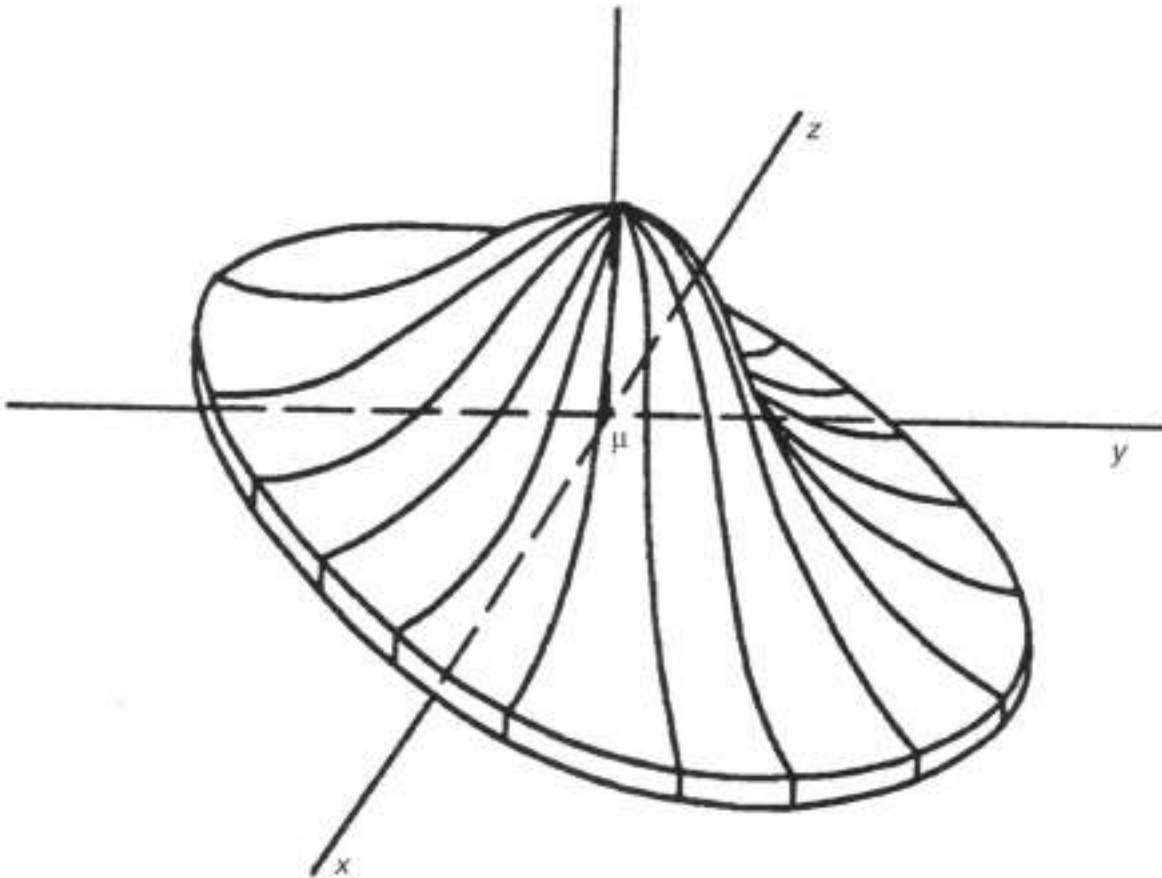
Bias in the Survey Instrument. Survey results can be biased if the instruments (the questions) are loaded with hidden values that elicit certain reactions from the respondents. During the early days of the Soviet Union's dissolution, the government of President Mikhail Gorbachev objected to a referendum in a breakaway republic that asked, "Would you like to be free, or are you willing to be dominated by the forces of communist Russia?" Naturally, the election results were never in doubt. In public policy analysis, a faulty survey instrument can often distort the results. For example, when asked in a general context, most people are likely to show their preference for such public goods as a clean environment, abundant wildlife, or a safer and well-maintained infrastructure. However, these answers may not truly reflect public opinion, because people are just as likely to change their minds when these same questions are asked with reference to how much the programs may cost them personally.

Furthermore, calls for subjective judgment can create biases. Asking people to categorize something as "high," "low," "large," "small," and so forth, elicits the subjective nature of people's judgment, which may stand in the way of determining the "true" opinion of the public.

Random Sampling Errors

Random sampling errors occur because of unexplained variations around "true" population values. Because these variations are random (or are caused by factors that cannot be determined), they can fall in any direction around the true values. For example, if you throw stones at the middle of a target, some will hit the bull's-eye and some will not. The stones will have a strong central tendency because you have aimed at the bull's-eye. Therefore, however poor a marksman you are, the stones will form a more or less normal distribution (unless, of course, your arm gets tired and creates a bias as a result), which will look like a Mexican sombrero. Similarly, if the errors are randomly distributed around the mean in a three-dimensional situation, they will form normal distributions in any direction around the mean (μ). I have attempted to depict this in Figure 8.3. As you can see, the true population value (of, say, the proportion of the population favoring handgun control) is μ . However, because of the presence of random variations, the sample value may turn out to be different from this true value.

Because these variations are random, we cannot completely eliminate them. Instead, we try to increase the precision of the sample results by reducing the random error factor. How can we do this? Recalling the discussion of the law of large numbers in chapter 7, we know that as the size of the sample increases, the domain of the error term shrinks. So when the number of samples approaches infinity or the sample size becomes the same as the entire population (in which case it is no longer a sample but a complete enumeration), the sample value becomes equal to the true population value. As already mentioned, however, costs often preclude a complete enumeration. Hence, we need to design a survey that is within our financial means yet provides us

FIGURE 8.3 Random Sampling Errors

with a tolerably precise result—that is, one with a reasonably low error level. We will discuss this more at length when we tackle the question of sample size later in this chapter.

Armed with the knowledge of possible pitfalls, we may now discuss the process of designing a survey. This process essentially centers on four questions: (1) Who is chosen? (2) How many are chosen? (3) Which questions should be asked? and (4) What is the mode of polling?

Choosing the Sample Population

There are numerous methods of choosing a group of subjects, some of which are called **objective sampling** or **probability-based sampling**, in which the sample subjects are drawn according to a set of rules meant to maximize the probability of achieving an accurate account of the population. In contrast, one may use the **judgmental method of sampling**, in which one uses specialized knowledge about the population to get to the correct information.

The following are methods of objective sampling.

- **Random sampling.** When we conduct random sampling, each element in the population has an equal chance of being picked. There are many ways of ensuring that the samples are drawn in a truly random fashion. For example, you may use

computer-generated random numbers to pick the subjects. Or, in a large metropolitan city, you may go by the listings in the telephone book (say, by choosing the tenth person on each page). This method is most likely to give us the desired results when the population is relatively homogeneous. If we are interested in gathering information on social norms and mores, a random sample may be the most appropriate design, because each member of the society has an equal voice when determining individual attitudes.

- **Stratified sampling.** In contrast, when a population is divided into distinct subgroups, or *strata*, it may be better to identify each stratum of society and poll it separately. An example may clarify the situation. Suppose a school district is considering switching from the traditional academic calendar to year-round schooling. To assess the reaction to this proposed change, the school district wants to poll the residents of the district. Because the population is fragmented along the lines of income and race—and because the district authorities want to be sensitive to the special needs of single parents—it is advantageous to poll each subgroup separately.

- **Cluster sampling.** When an extremely large population must be polled, a pollster may find it difficult (or too expensive) to include each member on the list of potential subjects from which the sample is to be drawn. Instead, the pollster may concentrate on a microcosm of the larger population. Let us suppose that we need to examine an aspect of urban America, and we find it impossible to reach every city. We may argue, then, that a particular city (say, Buffalo, New York) can serve as the “typical city.” Hence, we can poll the residents of Buffalo and draw conclusions about the rest of the urban areas of the United States. As you can see, the accuracy of this procedure depends on how representative of the population the cluster is.

Choosing the Size of the Sample

The development of theoretical statistics has given us some of the most powerful tools of quantitative decision making. One of these tools helps us determine sample size by linking the concepts of confidence interval and sample error. As discussed in chapter 7, the confidence interval is calculated using the following formula:

$$-Z_{\alpha} \times \frac{\sigma}{\sqrt{n}} > \bar{X} - \mu > +Z_{\alpha} \times \frac{\sigma}{\sqrt{n}} \quad (8.1)$$

where Z_{α} is the value from the Z-score table (Appendix C) at the desired level of confidence, α ; σ is the population standard deviation; n is the population size; $\bar{X}$ is the sample mean; and μ is the population mean.

This formula tells us that we can be confident at any desired level of α that the difference between the sample mean and the population mean will fall within a particular band. We can rewrite this formula as

$$(\bar{X} - \mu) \pm Z_{\alpha} \times \frac{\sigma}{\sqrt{n}} \quad (8.1a)$$

or

$$\mu = \bar{X} \pm \left(Z_{\alpha} \times \frac{\sigma}{\sqrt{n}} \right) \quad (8.1b)$$

The problem with using this formula is that it uses the population standard deviation, which is unobservable. We can estimate the population standard deviation from the sample standard deviation by correcting the formula with the degrees of freedom:

$$\mu = \bar{X} \pm Z_{\alpha} \times \frac{s}{\sqrt{n-1}} \quad (8.1c)$$

where s is the sample standard deviation and n is the size of the sample population.

Let us take an example. Suppose that from a sample of 100 residents of a city, we learn that 67 percent support a proposed referendum. We also know that the survey has a standard deviation of 0.15. To be safe, we want 95 percent accuracy. Therefore, we can consult the Z table for the value of $Z_{.05}$, which turns out to be approximately 1.96. Using these numbers in equation 8.1, we can be 95 percent certain that the population mean, which we cannot observe, will fall within the range

$$\mu = \pm 1.96 \times \frac{0.15}{\sqrt{100-1}} = 0.67 \pm 0.0295$$

That is, we can be 95 percent sure that the opinion of the population will vary within the range of approximately 70 percent and 64 percent. In other words, the sample result contains a sample error of ± 3 percent (by rounding off 0.0295).

As you can see, there is clearly a trade-off between the need for accuracy and its cost. If you want a very low sample error, you will have to interview a lot more people, which increases the cost of the survey. For example, you need a high degree of accuracy to predict the outcome of a close election, but you might not demand such a high standard for other policy-related polls.

Choosing Effective Survey Instruments

Although choosing the size of the sample is mechanical, the real art (derived from experience and creativity) of conducting a survey lies in designing the questionnaire. Many concerns arise when developing a questionnaire; we may classify them in five broad categories:

- *Be on a level with the respondents.* Questionnaires must be developed with the respondents in mind. If you are developing a questionnaire for middle school children, would you ask questions with words that are beyond their vocabulary? Certainly not. And since you are asking youngsters, you would be careful about the way you phrase your questions. The same caution applies to questions put to adults. The question "Do you belong to the poor, middle, or wealthy class?" would draw a mixed reaction, because definitions of these class boundaries vary considerably. Instead, you can get a much more specific answer by asking, "Which of the following income groups do you belong to—less than \$10,000; between \$10,001 and \$20,000; . . . ?"

- *Choose the right format for answers.* Questions in a survey can be either open-ended or fixed. With an open-ended question, answers can be long and descriptive. To determine positions on handgun registration, questions may be phrased as, "What is

your position on handgun control?" Alternatively, questions may be put within a fixed format, such as

"I believe all handguns within the state should be registered."

____ Strongly Agree

____ Agree

____ Neither Agree nor Disagree

____ Disagree

____ Strongly Disagree

____ No Opinion

Both types of questions have advantages and drawbacks. Open-ended questions draw out a full spectrum of responses without restricting respondents. However, the answers to these questions, by virtue of being unstructured, cannot be easily fitted into a specific coding system for quantitative analysis. A highly structured question, however, may miss an answer that many respondents might have chosen.

- *Avoid words that provoke emotion.* If your goal is to develop an objective set of questions, you should avoid terms that evoke strong emotional reactions. For example, you are trying to determine public support for the U.S. government's food aid to drought-stricken parts of Africa. Consider three questions:

1. "Do you support U.S. aid to drought-stricken parts of Africa?"
2. "Do you support U.S. aid to the starving people of Africa?"
3. "In view of all the problems facing this country, do you support U.S. aid to Africa?"

Although in general, few people are against food aid to a starving group of people, it is fairly obvious that the answers to these similar questions are likely to be quite different. However, to an unsuspecting audience, you can report the results of this survey as "people's opinion on U.S. aid to Africa" by picking the one that suits your purpose the best.

- *Choose the proper sequence of questions.* The sequence in which questions are put to respondents can be crucial. The questions can be asked in a random sequence, or they can be asked in a funnel or an inverted funnel sequence. If respondents are likely to be familiar with the broad issues, it is better to use a funnel-like sequence, starting with the most general question and then narrowing down to a more specific one. Suppose a school district is debating the distribution of literature on birth control devices to its high school students. Because most people hold some kind of opinion on the state of family values, you can start by asking respondents to assess the state of family values and then, within that context, ask their opinion about birth control. In contrast, most people are not well informed about the problem of the external trade deficit and may require some help in framing the issue. Therefore, in such a survey, you may do well by starting out with a set of specific questions, which can help them focus on the broader issue.

- *Choose the appropriate length of questionnaire.* Because conducting a survey is expensive and time-consuming, there is a tendency to load questionnaires with as many questions as possible. This tendency is to be resisted, as most people lose interest in answering questions if they are too long or there are too many. This is particularly true for mail or telephone surveys. Greater latitude is often possible when interviewing face to face.

Choosing Polling Methods

Surveys are about asking questions. Determining the mode of questioning is an extremely important aspect of survey design. Much thought must be given to choosing an appropriate medium. Generally speaking, there are three methods of interviewing:

- *Face-to-face interviews.* The personal, face-to-face interview is the most expensive, although probably the most accurate, method of gathering information about a population. Face-to-face interviews are particularly appropriate when a questionnaire has many open-ended questions or is lengthy. Frequently this form of personal interview can put subjects at ease with the interviewer. Yet this personal contact may have its drawbacks, as the interviewer and the interviewee may find each other less than acceptable, and as a result, the survey results may be skewed.

- *Telephone interviews.* Telephone surveys may be the least expensive type of survey to administer and are not likely to create systematic biases. They are most effective in conducting quick, rather than in-depth, surveys. However, people are often squeamish about opening up to a faceless voice over the telephone and tend to be guarded in their responses.

- *Mailed questionnaires.* Relatively inexpensive and free of the influences of individual interviewers, mailed questionnaires allow subjects to think over their responses and can also be used to elicit long answers to open-ended questions. Also, once the questionnaires are returned, they can be processed and data can be recorded, tabulated, tested, and reported with relative ease. In many cases these surveys are recorded on paper and scanned by a computer, creating files for subsequent testing. However, mailed questionnaires often suffer from extremely low response rates. While for personal interviews response rates are typically about 95 percent, for mail surveys these numbers have been observed to be between 2 percent and 40 percent. A low response rate should warn us that unless we are careful, the results can exhibit a self-selection bias. In other words, as discussed earlier, if the rate of returns is extremely low, those who have taken the initiative to fill out the questionnaires and mail them may represent a special subset of the population.

QUANTIFICATION OF SURVEY DATA

Survey results are typically expressed in absolute numbers or in percentages and ratios. We are also told survey results in terms of means. Because it is difficult to report the results of open-ended questions in numbers, most surveys use fixed-format answers. As noted at the beginning of the chapter, certain questions are answered in numbers (income, age, years of education, and so on), others in nominal scales (sex, race, religion,

and the like). Finally, data that relate to matters of attitudes and values cannot be readily translated into numbers. To express these variables in numbers, we use predetermined scales.

In the social sciences, **Likert scales** are the most commonly used. With a Likert scale, respondents are asked to express their feelings on a continuum varying from highly positive to highly negative. On a five-point scale, the respondents will be asked whether they “strongly agree,” “agree,” are “undecided,” “disagree,” or “strongly disagree” with a particular statement. These answers can then be assigned weights either in ascending (1, 2, 3, 4, 5) or descending (5, 4, 3, 2, 1) order. On the basis of their answers, respondents will be given points, which will indicate their individual stances on particular issues, or the average of the responses can be calculated for determining the “average” attitude or feeling.

REPORTING SURVEY RESULTS

Surveys are conducted to gather information and to test a specific set of hypotheses. It is very important, then, to report the results of a survey properly. Most are reported with tables, graphs, and charts. I will discuss these techniques in chapter 9. However, because survey data can be misused and misinterpreted, you must spell out in detail the assumptions, sampling method, sample error, level of confidence, and often the questions themselves. In a typical example, the *Los Angeles Times*, when reporting the results of a nationwide opinion survey, informs readers how the poll was taken:

The *Times* poll interviewed 1,146 registered voters nationwide, by telephone, from August 12 to 14. Telephone numbers were chosen from a list of all exchanges in the country. Random-digit dialing techniques were used to ensure that both listed and nonlisted numbers had an opportunity to be surveyed. Results were weighted slightly to conform with census figures for sex, race, age, education, and household size. The margin of sampling error for the total sample is plus or minus three percentage points. For certain subgroups, the error margin is slightly higher. Poll results can also be affected by other factors, such as question wording and the order in which questions are presented.⁶

CONDUCTING FOCUS GROUPS

Focus groups are becoming increasingly popular as a tool for gauging public opinion regarding policy issues at all levels of government. For a focus group, a small representative group of the target population, usually numbering between eight and twelve, is assembled for in-depth discussions on a certain topic. Under the guidance of a trained interviewer, the group discusses structured questions relating to the central topic. Focus groups can shed light on complex areas of public policies and other important issues that may be missed by a survey, which can be too structured (see “A Case in Point”).

In conducting focus groups, which can be powerful tools of policy formulation, you must address a number of methodological and procedural questions. Unlike a survey, there is no set methodology for focus groups. Since they are small and the questions open-ended, strict statistical techniques cannot be used for their verification. A great



A CASE IN POINT

Focus Group for Growth Management Policies

The problem of managing economic growth confronts many regions of the nation. Palm Beach County is one of the largest counties in Florida, comprising 2,223 square miles and thirty-nine municipalities. In the 1990s, the county experienced some of the fastest economic growth in the United States. Located between the ecologically sensitive Everglades, Lake Okeechobee, and the Atlantic Ocean, the county's population grew by an incredible 20 percent per year. In 1998 the Sierra Club ranked Palm Beach County as one of the most sprawling medium-sized metropolitan areas in the nation. Therefore, the problem of managing runaway growth is paramount to the region's planners. However, this issue is quite complex, since it aims to reconcile often contradictory interests of the citizens.

In 1997 the county introduced its Managed Growth Tier System, which divided it into five distinct regions: urban-suburban, exurban, rural, agricultural reserve, and the glade. Palm Beach County's plans were tested by more than twenty-two focus groups for a year-and-a-half. These groups, representing various interest groups in the region, discussed the plans to fine-tune various proposals. At the end, when the growth management plan was approved, it received the American Planning Association's award for 2000.⁷

Discussion Points

1. When is it appropriate to use focus groups in public policy analysis?
2. Why was Palm Beach County successful in using focus group findings?
3. Consider a particular issue in your community. Do you think you can recommend the use of focus groups to determine citizens' positions on the issue? Why or why not?

deal of advanced thinking and planning must compensate for methods that are less than rigorous. I can make a few suggestions for developing successful focus groups:

1. *Determine a clear-cut set of goals.* For example, in Palm Beach County, it was decided to concentrate only on respondents' views on the proposed public policy of dividing the region into five distinct zones of growth.
2. *Identify the critical characteristics of potential participants.* In the Palm Beach County case, do the focus groups include "constituents," meaning a group drawn from the resident population, or is the purpose to understand the views of key interest groups (for example, the Chamber of Commerce, small business groups, commuters, builders and developers, agricultural workers, environmentalists, and so forth)?
3. *Recruit skillful group leaders.* You should note that it is extremely important to use skillful discussion leaders who can keep the focus of the group tightly on the

issues at hand. Discussion leaders must balance the need for the participants to freely express their opinions on controversial issues with that of avoiding futile arguments.

4. *Analyze the discussion for summary and conclusion.* There is no set procedure for analyzing focus group discussions. They can sometimes be taped with the permission of the participants, or the discussion leader may take copious notes for analysis. The summary and conclusion must be presented in a clear and concise manner so they answer the decision makers' pertinent questions.
5. *Determine the number of focus groups.* Because focus groups are typically composed of eight to twelve people, you need to have a number of such groups to get a feel for the community's diverse views on public policy questions.

SECONDARY DATA

Opinion polls and focus groups are excellent tools for needs assessment or for prioritizing needs for a local government. However, most public policy research involves the use of secondary information, collected by someone other than the researcher. When you are confronted with an unfamiliar topic, you may be at a loss as to where to begin. Chances are there is a vast amount of information, but you must be able to conduct an effective search so that you can find the information you need.

Let us take an example. Suppose you are a junior analyst, working for your city, which wants to address a great deal of abandoned or dilapidated housing in the downtown area. The city wants to provide an incentive to homeowners to make substantial improvements through a proposed property tax rebate program. Under this program, homeowners in the designated area will qualify for a property tax rebate for three years if they improve their properties. Your job is to research the project's effectiveness.

To fully understand the issues relating to your project, you must first find out what work has already been done on such issues. Your literature review will help you

- gain a fuller view of what other researchers have discovered through their investigations, as well as point to areas that still need to be addressed;
- identify causal linkages that may help explain your hypothesis;
- conceptualize and measure key concepts; and
- identify data sources.

Thus, a literature review will quickly tell you about the experiences of other cities that have experimented with similar programs. These published books and articles will help you develop your own hypotheses and methods of formulating the analysis. On the basis of the literature, you may decide to try to find out if a tax rebate does in fact spur investment in housing renovation.

However, the decision to improve one's property is complex and depends on factors other than property taxes. It involves the calculation of, for example, the condition of the neighborhood (neighborhoods with higher property values would provide investors with higher returns), levels of property tax rates, and current economic conditions. A review of the relevant literature will help you identify the factors that are important to a potential investor in property improvements.

Once you know these causal linkages, you can start thinking about how to define and operationalize (numerically measure) these variables. You can operationalize the economic condition of a neighborhood either by its average household income or by the median property value. Although these two figures are closely linked, they do not convey an identical measure of the neighborhood's relative prosperity. A quick reading of the literature will provide you with a good idea of which measures other researchers have used and why.

After you have decided how to operationalize the relevant variables, you need to find data containing such information. For example, if you know that fifty cities have attempted tax rebate programs in the past, you may start searching for comparable data on those cities. Many sources can provide you with the information you seek (see this and following pages). Once you have gathered the relevant data, you may proceed with your analysis.

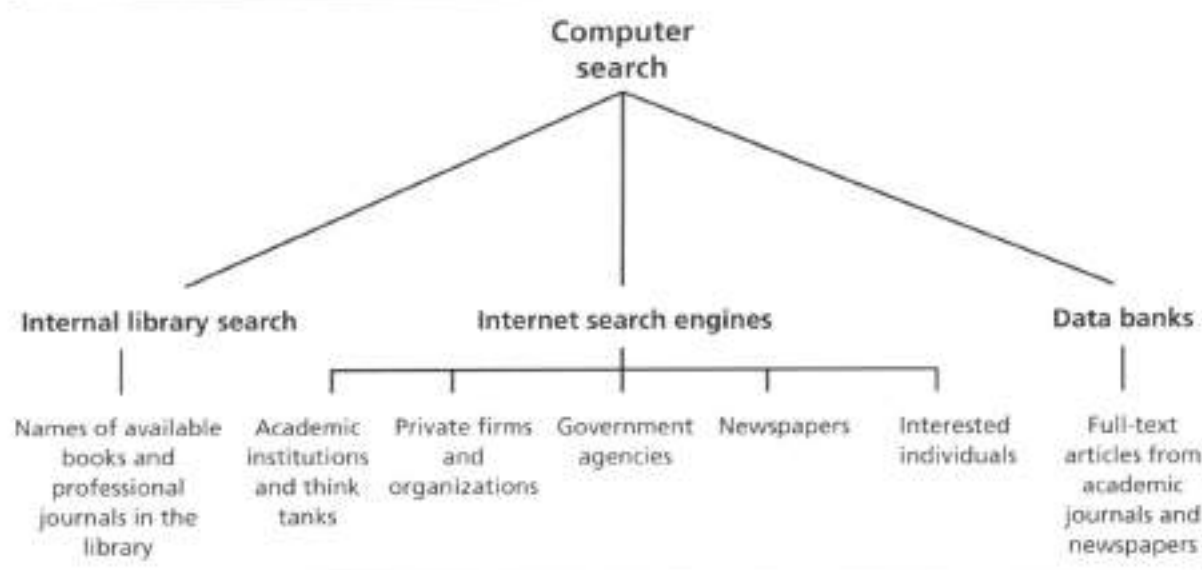
Searching for Information

Thanks to advances in technology, sources of secondary information are unlimited, and most college and university libraries offer their students online services. However, before you start searching—and perhaps get frustrated—you must develop a strategy. Not all the searches you attempt will deliver the right information. I have classified information sources into three broad categories (see Figure 8.4). You can conduct an internal search within a library for books and scholarly journals. During such a search, you can explore by using an author's name, a title, a subject, or words in the title. If you know the names of prominent authors in the field, or you know exact titles, you can easily find out if the library has sources you would find useful. If you do not have a good idea about the topic, it is always smart to start with a recognized textbook, because it provides you with a well-rounded introduction to the subject matter. You can also search by key words. For example, for our inquiry, I typed in "property tax" and found a number of useful books. Once you come across a relevant book or article, look at the list of references. These citations can lead you to other relevant works. This process of building a bibliography is called **pyramiding citations**, because you can develop a substantial list from a small number of initial sources. In a library search, you will have to evaluate books and journal articles to see which of these will be useful to you.

Online Searches

An Internet search, in contrast, offers a great deal of information without your having to leave your desk. In an online search, you quickly discover that a world of information awaits you. The problem, however, is how to focus your search to find pertinent information, and once you find that information, how to determine its reliability. In this intimately interconnected information network, the first problem is to find a proper starting point.

Today, most people are familiar with general search engines or portals, such as Google and Yahoo. These search engines are excellent tools for locating institutions, government agencies, particular newspapers, and people. If you Google "property tax," you will receive a large number of hits, which include, among others, tax consultants, the Internal Revenue Service, individuals trying to sell books on how to lower your

FIGURE 8.4 Searching with the Computer

property tax bill, lists of unclaimed properties, and state property tax forms. Conducting an internal search on an institution's Web site also can be useful. For example, on the Web site of the Brookings Institution, a leading think tank, a search for "property tax" generates a list of the institution's publications on the subject. Another strategy is to search in the online archives of a news source.

All search engines depend on key word searches for accessing the right information. As stated above, this type of search (often called a Boolean search) allows you to specify a word or phrase that best represents your area of inquiry. Some engines allow you to use the operators AND, OR, and NOT to create either a very broad or a very narrow search. Say you are interested in looking for articles on school district consolidation. You began the search with the broad key word "school financing." Because this is too broad a search, you get a large number of hits, most of which are of little or no interest to you. In that case, you can narrow the search by adding another key word following the operator AND, which combines search terms so that each result must contain both terms, such as "school AND consolidation." If you are interested in further narrowing the search, you could add another word (say, "Midwest" or "Kansas") along with the two, separated by another AND operator. If you want to broaden your search, use the operator OR. The OR operator combines search terms so that each result contains at least one of the terms. For example, "school OR consolidation" will give you results that contain either term. Some advanced searches also allow you to use the operator NOT. This operator excludes the results that contain the term that follows NOT. For example, if you are interested in the consolidation of individual schools, not school districts, you may restrict your search to find results that contain the term "school" but not the term "district."

The results from a search often hide the information you need within a heap of unrelated, inaccurate, or incomplete materials and links. It is up to you, then, to determine whether a site is to be trusted, if the information is accurate, or if a more advanced search would produce better results. A search for the name Einstein gives me information on the

life and work of the famous physicist, but it also produces sites for a garage, a laboratory auction house, and a chain of bagel stores. It is therefore important to remember that anyone can put up a Web site. It is also important to remember that unlike a published book, postings on the Internet have no physical existence. Therefore, a site that looks interesting and useful to you now may not be there tomorrow, or it may change significantly. This lack of permanence can cause problems when you include references to Web sites in your report. There is no better example of this problem than Wikipedia, the online encyclopedia that anyone can edit. The open-source nature of Wikipedia makes it a very up-to-date resource, but it also creates a problem: there is no certainty of the accuracy of its information. Therefore, one must be extremely careful in using Wikipedia and should crosscheck all information from other, more reliable sources. Furthermore, Wikipedia tends to be more accurate in reporting specific facts rather than providing information on issues that can evoke emotion or subjective judgment. You are more likely to find reliable information on the date of birth of Napoleon Bonaparte than on the death penalty, abortion, or gun control.

So exactly how do you judge the reliability of a Web site's information? The answer to this question is less complicated than it may at first appear. Because of the wide variation in the reliability of printed sources as well, you should not accept any kind of information without a critical review. You should always test the information's validity, accuracy, and authority. When getting information from the Web, see if the information follows a valid sequence of reasoning. Analysis based on critical reasoning differs from propaganda based on rhetoric only. However, even if a piece appears to be based on sound reasoning backed by empirical observations, you may question the validity of the data. If the data are not widely available, you may have to establish their accuracy through cross-referencing. That is, you may want to know if the data have been used in published work in a reputable scholarly journal, magazine, or newspaper. You are much safer in accepting the validity of the information if the source is a reputable organization, such as a university, a well-known independent research outfit, a well-regarded publishing house, or an official government entity. But even then you should be aware of the possible shortcomings of the information. The data may contain deliberate biases stemming from ideological positions or from vested interests. If I am looking for information on the effects of environmental pollution, I may have reason to be skeptical of data provided either by an environmental activist group such as the Sierra Club or by an industry with a stake in the outcome of the public debate. In such a situation, you should look for corroboration by other independent sources.

If you are interested in articles published in scholarly journals, magazines, or newspapers, you can go to specialized databases. Databases are different from general search engines in that they cover only a specific pool of information. Some may cover a set number of professional journals. You can log on to such a database and look for articles published in these journals. Most of these databases allow users to download entire articles without having to go to a library. Searching for "property tax" and "abatement" in the ProQuest database, I found some excellent articles detailing the experiences of Syracuse, New York, and Dayton, Ohio, with regard to tax rebates for property improvements.⁸ If you are interested in publications from the top social science journals, you should search JSTOR. You may also find many other excellent articles in the area of your interest in various other databases. I discuss a few of these later in these pages. However,

I should mention that accessing these databases can be an expensive proposition, since the information is proprietary. However, if your university has paid the required fees, you may be able to access these databases as a student.

For public policy analysts, perhaps the best sources of secondary data are U.S. government publications (see “A Case in Point”). Appendix A includes an example of one of these publications—a report conducted by the U.S. Government Accountability Office (GAO) on the state of poverty in America. Government agencies in the United States are inveterate collectors of statistical information. Furthermore, most government organizations are part of this enormous information-sharing network. Although a multitude of extensive databases are available and new ones are being created, I mention a few important ones at this point:

- CIAO (Columbia International Affairs Online) (www.ciaonet.org) is an excellent source of policy-related papers. In addition to a number of scholarly journals, it provides policy briefs and working papers from a large number of think tanks and research institutes.
- EBSCOhost Academic Search Elite (www.ebscohost.com) provides access to a large number of journal articles from various academic disciplines. It covers more than 1,200 periodicals starting around 1990.
- LexisNexis Academic (www.lexisnexis.com/us/lnacademic/) allows you to retrieve full-text versions of national and international news articles. Its coverage generally starts from the late 1980s.
- ProQuest Research Library (www.proquest.com/en-US/catalogs/databases/detail/pq_research_library.shtml) is one of the most comprehensive sources of scholarly articles. Its core research module covers traditional academic journals and is supplemented by subject-specific modules, such as Arts, Education, Humanities, and International and Multicultural Studies. Coverage begins in 1986.
- PAIS International (Public Affairs Information Service) (www.csa.com/factsheets/pais-set-c.php) is also one of the most comprehensive data sources, containing an index to journal articles. This database is particularly important to us because of its primary focus on public policy. Its coverage starts in 1972.
- JSTOR (www.jstor.org) provides access to full-text scholarly journals in the humanities, sciences, and social sciences and some in the natural sciences and population studies. This database provides full texts of published articles in the most prestigious scholarly journals in the fields of social science, demography, environmental sciences, and statistics.

Having gone through the relevant literature, you are now ready to collect numerical data on your topic. You should start with the annual *Statistical Abstract of the United States*, the annual *Economic Report of the President*, the *Handbook of Labor Statistics*, and the *U.S. Census Bureau Catalog and Guide*. Several excellent publications provide state and local government data; the *Census of Governments*, the *County and City Data Book*, the *County Yearbook*, and the *Municipal Yearbook* are a few of them. There are also excellent sources for international data. However, international data are often not comparable, as different governments use different criteria for classifying information, which makes cross-national comparisons problematic. Nevertheless, a number of large international agencies compile information cross-nationally. For example, if you are interested in basic economic, political, and demographic information, you may start



A CASE IN POINT

Politics and Government Statistics

When we use data from U.S. government publications, from census figures to the rate of inflation, we tend to take their objectivity for granted. Yet, like many activities of government, its data collection is not completely free of political considerations. William Alonso and Paul Starr point out that “official statistics do not merely hold a mirror to reality. They reflect presuppositions and theories about the nature of society. . . . Political judgments are implicit in the choice of what to measure, how to measure it, and how to present and interpret the results.”⁹ This political intrusion can cause serious public policy implications in the collection and interpretation of the data.

Stephen Feingold provides a vivid example of how politics affected the effort of the Centers for Disease Control and Prevention (CDC) to collect data on AIDS.¹⁰ The problem of defining the disease with precision complicated the collection of data. A relatively large number of people carry the human retrovirus known as HIV. Of those who do carry it, only a small portion actually develops AIDS-related complex, and a much smaller proportion of this group actually develops full-blown AIDS. The population groups with the highest incidence of AIDS are homosexual men, intravenous drug users (and their children), and hemophiliacs.

In the 1980s, data on the spread of AIDS were based on an antiquated system of counting that used a much more stringent definition of the disease. Because of its political position that the spread of AIDS was not a serious matter for most Americans, the administration of Ronald Reagan was reluctant to allocate money for a more complete survey.¹¹ Further complicating the matter was the conflict between the measurement issue and both the government’s “war on drugs” and a general prejudice against homosexuals. When the CDC proposed an intensive national survey for an accurate count, Senator Jesse Helms, R-N.C., argued that it was not intended to stop the spread of AIDS but to provide evidence that homosexuality was “just another normal lifestyle.” In the end, the survey was conducted in 1992, funded by private donations. When the results came out, much to the surprise of the scientific community, the threat of a rapid spread of the dreaded disease was found to be a lot less than expected.

Discussion Points

1. Discuss how politics can influence data gathering by even the most trusted government agencies.
2. As a policy analyst, how can you guard against biases in the information used in your analysis?

with the Central Intelligence Agency’s *World Factbook* (<https://www.cia.gov/library/publications/the-world-factbook/>). You may also consult the World Bank (www.worldbank.org), the *United Nations Statistical Yearbook* (<http://unstats.un.org/unsd/syb/>), the *International Trade Statistics Yearbook* (<http://unstats.un.org/unsd/trade/default.htm>), and the United Nations National Accounts Statistics (<http://unstats.un.org/unsd/snaama/Introduction.asp>).



A CASE IN POINT

The Politics of Data Collection: The Case of Terrorism Research

In the ideal policy world, hypotheses are tested with empirical evidence. By analyzing data, we either accept or reject hypotheses and project the future outcomes of various alternatives. For instance, when Congress proposes a new tax, it can ask the Congressional Budget Office (CBO) to project the impact of the tax on the economy. The CBO uses readily available, detailed macroeconomic data in its statistical models to get the required answers. However, in many other areas of policy analysis, our research agendas can be significantly hamstrung by the paucity of empirical information. As a result, policymakers may be forced to make decisions on the basis of educated guesses, conjectures, or simply conventional wisdom. Because of the “tyranny of data,” the availability of statistical information can influence the research agenda.

At the height of the Cold War, when the United States was worried about the causes of open hostilities between nations—notably, between the two superpowers—funding started to flow to research projects on wars. University of Michigan political scientist David Singer and historian Melvin Small began collecting data on wars under the title *Correlates of War*.¹² The emphasis shifted in the late 1960s, when the United States and Western Europe were rocked by acts of political violence by disgruntled youth. This new reality forced the government to sponsor data collection on a different kind of social conflict. Many of these data sets were published, but they were collected at the national level, so they did not offer researchers a finer grain to analyze violence at regional levels or by groups.¹³ However, as the riots, assassinations, and coups d'état that rocked the world in the 1960s and early 1970s subsided, the funding for these data collection efforts quickly dried up.

The availability of data was altered once again after a huge truck bomb devastated the Alfred P. Murrah Federal Building in Oklahoma City in 1995, taking 168 lives in the process. The bomber turned out to be homegrown: a decorated former U.S. army soldier, Timothy McVeigh. In response, a nonprofit organization, the Memorial Institute for the Prevention of Terrorism (MIPT), was established (it later received help from the U.S. Department of Homeland Security) and began collecting and posting terrorism data on its Web site. The MIPT's data sets were managed by the RAND Corporation, but after several years, its funding became depleted, and the RAND Corporation took them off the Web.

In addition to these vast amounts of economic, political, and social information, a huge array of data on attitudes and opinions are available. The Institute for Social Research at the University of Michigan collects and compiles data on many aspects of social, political, and psychological attitudes.

Finally, if your research project involves a particular government agency, the agency's annual budget reports can supply a great deal of information. The agency may collect for its own internal use a good deal of information that may not be available in a ready-to-use form. However, you may be able to make effective use of these pieces of

Horrific attacks by the members of a small group who plotted from the dusty caves of Afghanistan shocked the United States on a September morning in 2001. Stunned, politicians, pundits, and the general public started groping for answers about why it happened and how to prevent future attacks. Those who endeavored to find answers, however, found themselves in a quandary caused by the lack of reliable data on the subject. As a result, everyone came up with his or her own theory. President George W. Bush proclaimed that “they hate us for our freedom.” Others, notably in the liberal camp, assumed that poverty caused these men to hate Americans more than they loved their own lives. As a startled nation woke up to the freshly realized threat, there were calls for better collection of data.

This call spurred a debate over who should collect data on terrorism. Terrorism data are collected by governmental agencies such as the U.S. State Department, intelligence agencies, for-profit groups, think tanks, and university-affiliated entities. But because *terrorism* is an intensely political term and defies a single definition, any agency that collects data will be suspected of injecting its own ideological orientation into it.¹⁴ The State Department, for instance, was roundly criticized for politicizing the data that it started to report soon after the 9/11 attacks.¹⁵ As a result, it stopped publishing data. The intelligence agencies in every country collect data, but they are classified and therefore not open to scrutiny by the scholarly community. For-profit organizations, such as ITERATE (International Terrorism: Attributes of Terrorist Events), offer information—sometimes at a stiff price—to private firms, academic institutions, and individual researchers.¹⁶ Since terrorism is a social phenomenon, its wide analysis and understanding is crucial for constructive policy debates. Therefore, the Department of Homeland Security established the National Consortium for the Study of Terrorism and Responses to Terror (START), located at the University of Maryland. Funded by the U.S. government and supervised by independent scholars, START offers the best source of information on global terrorism. Yet in recalling the history of government funding, it is conceivable that if there are no sensational terrorist attacks directed at the United States, this worthwhile effort may also be abandoned as public attention shifts elsewhere.

Discussion Points

1. Why is it important to know who should collect data on terrorism?
2. How can we ensure objectivity in data collection?
3. How do sensational events affect how data are collected?

information by arranging them in a proper way. Most of these sources are open to public use and are available on the agencies' Web sites. If you are looking for national-level or local-level data, you may consult the following resources:

- The *Statistical Abstract of the United States* (www.census.gov/compendia/statab) provides a large number of national census data sets. It also contains most comprehensive data sets on state and local governments. For more detailed sets, you may have to purchase the CD-ROM.

- If you are interested in data on the Consumer Price Index and information related to labor, visit the Web site of the U.S. Bureau of Labor Statistics (www.bls.gov).
- If your library subscribes, you may also access the online data bank of STAT-USA. This data bank allows you to access daily economic news, frequently requested statistical releases, the National Trade Data Bank, the Economic Bulletin Board, and the U.S. Bureau of Economic Analysis.

Searching the Old-fashioned Way

If your library does not allow you to search electronically, the traditional, time-honored places to start are the published versions of the *PAIS Bulletin*, the *Social Sciences Citation Index*, or the *ABC Index of Political Science and Government*. If you are specifically interested in urban issues, there are selective publications, such as the *Index to Current Urban Documents*. Many specialized indexes for various areas of the social sciences, public policy, and planning are available, as are subject-specific abstracts of important articles in the fields of business, economics, education, law, political science, public administration, public policy, sociology, urban and regional planning, and so on.

Finally, for a thorough literature survey, you may also build on other students' accepted work. This is an often overlooked area of a literature search. Although most Ph.D. dissertations and master's theses are not formally published, upon their acceptance for degrees, they become public documents. Because the students who wrote them have already spent a great deal of time developing bibliographies, access to relevant theses or dissertations can be extremely useful to you as a researcher, especially if you are not familiar with the topic. Accepted dissertations and theses are available online from *Dissertation Abstracts/Digital Dissertations*, a proprietary database to which many libraries across the nation subscribe.

When Data Are Not Available

Having gone through numerous data sources, you might feel frustrated. The specific information you are looking for does not seem to exist anywhere. It may be that your topic is too specific. Or it may be that for political or cultural reasons, such data have not been collected (see "A Case in Point"). You may not find time-series data on the homeless population in your city because the subject is too specific and no agency is currently collecting the information. In some cases, the data for a few crucial years may be missing from an otherwise complete series. Or, suppose you are interested in estimating the possible effect of a new convention center on your city's economy. You will not find much information, because the convention center has not yet been built. Suppose you are looking for data on smoking-related deaths on a cross-national basis. It may be that because of a lack of awareness, such data have not been collected for many countries around the world.

In such cases, your job becomes tenuous. You may consider inferring the data. That is, if you can find a comparable city with a similar demographic, cultural, and economic background, you may draw a parallel. These kinds of inferences are often permissible when no data are available. However, if you must have a series prepared through inference or interpolation, you must make absolutely clear to the reader the nature of your data and the procedure by which they were obtained.

KEY WORDS

Cluster sampling (p. 186)	Probability-based sampling (p. 185)
Focus groups (p. 190)	Pyramiding citations (p. 193)
Interval scale (p. 178)	Random sampling (p. 185)
Judgmental method of sampling (p. 185)	Ratio scale (p. 178)
Likert scales (p. 190)	Revealed preference (p. 184)
Nominal scale (p. 177)	Sample population (p. 180)
Objective sampling (p. 185)	Survey design (p. 180)
Ordinal scale (p. 178)	Sampling biases (p. 180)
Population (p. 180)	Secondary data (p. 177)
Primary data (p. 177)	Stratified sampling (p. 186)

EXERCISES

1. There are many different ways to collect data on and measure poverty. The GAO report on poverty in America in Appendix A provides some examples of different measures. Why do you think the GAO analysts chose to use these measures? How did they go about collecting data? Why is it difficult to devise a universally acceptable measure for poverty?
2. What are primary and secondary data? Discuss the various scales of data. What are their relative strengths and weaknesses? What are their respective uses?
3. The city of Masters (population 150,000) wants to enlarge its airport to accommodate a growing need for a small commuter airline landing. The city wants to conduct a survey to assess public opinion on possible locations for the landing. Specifically, the city wants to know who would support and who would oppose such an expansion effort in certain already identified sites. The survey must be completed within three months. Design a survey, specifying the sampling method (random, stratified, and so forth), the number of people to be surveyed, questions to be asked, and the mode of interview.
4. Write a short essay on the biases of sampling. Collect information from the real world to elaborate on your points.
5. Consider the following three scenarios and give your recommendation regarding which type of sampling method (random, stratified, or cluster) or focus group to use.
 - a. The mass transit department in your city is considering a rate hike. Before it makes the final decision, it wants to learn more about the possible effects of such a hike on its customers.
 - b. The professional football team in your city is threatening to move to a different city unless yours is willing to spend many millions of dollars to expand and renovate the existing stadium. The city council wants to know the opinion of its citizens.
 - c. Your town is concerned about the scarcity of low-income housing. It is considering a new subsidy program that has been found to be effective in a different town of similar size. Your job is to design a survey to determine the effectiveness of the proposed program.
6. Suppose your town has received federal funding for educating the target population about sexually transmitted diseases. Your department has decided to conduct a survey. In an essay, discuss the problems you might face in obtaining a correct picture of the problem.
7. What are the advantages and disadvantages of conducting focus group research over a traditional survey design? Provide specific examples in which a survey would be more appropriate than a focus group and vice versa.

8. Consider an important issue facing your community. If you were to conduct a focus group, what steps would you take and what sorts of questions would you ask? Where would you meet?
9. Write a short essay on information gathering from the Internet. How would you ensure the credibility of the information?
10. Think of an appropriate focus group or survey design for your community. How would you estimate its cost?
11. Consider an important public policy issue. Prepare a bibliography and gather data through a Web search. If it is a controversial issue, such as gun control or abortion rights, rank the sites mentioned in your paper according to their reliability.

Notes

1. Professor Ted Gurr (Polity II database, Inter-University Consortium for Political and Social Research) quantified the degree of "democratization" of countries. Among several others, I have developed indexes of political instability. See Dipak K. Gupta, *The Economics of Political Violence: The Effect of Political Instability on Economic Growth* (New York: Praeger, 1990).
2. James Gleick, *Chaos: Making a New Science* (New York: Viking, 1987).
3. For a review of the implications of the Hawthorne experiment, see J. H. Smith, "The Enduring Legacy of Elton Mayo," *Human Relations* 51 (1998): 221–250.
4. Stanley Milgram, *Obedience to Authority: An Experimental View* (New York: Harper & Row, 1974).
5. For a detailed discussion of the human proclivity towards obedience to authority figures, see Dipak K. Gupta, *Path to Collective Madness: A Study in Social Order and Political Pathology* (New York: Greenwood, 2001).
6. *Los Angeles Times*, August 16, 1992, A6.
7. Stuart Meck, "Growing Smart: Initiatives and Applications—Palm Beach County Managed Growth Program," *Planning*, April 2000, 3–5.
8. On Syracuse, see William P. Barrett, "Willis Carrier's Ghost," *Forbes*, May 29, 2000; on Dayton, see Clifford A. Pearson, "Dayton Tackles Brownfields to Create Houses and Offices," *Architectural Record* 188 (June 2000).
9. William Alonso and Paul Starr, eds., *The Politics of Numbers* (New York: Russell Sage Foundation, 1987), 2.
10. Stephen E. Feingold, "Ethics, Objectivity, and Politics: Statistics in a Public Policy Perspective," in *Statistics and Public Policy*, ed. Bruce D. Spencer (Oxford, U.K.: Clarendon, 1997), 62–83.
11. *Ibid.*, 76.
12. For a description, history, bibliography on related projects, and data, see the Web site of the Correlates of War Project at www.correlatesofwar.org/cowhistory.htm.
13. See Arthur S. Banks, *Cross-Polity Time-Series Data* (Cambridge, Mass.: MIT Press, 1971); Charles Lewis Taylor and Michael C. Hudson, *World Handbook of Political and Social Indicators*, 2d ed. (New Haven, Conn.: Yale University Press, 1972); and Charles Lewis Taylor and David A. Jodice, *Annual Events Data* (Ann Arbor, Mich.: University of Michigan, Interuniversity Consortium, 1982).
14. See Bruce Hoffman, *Inside Terrorism* (New York: Columbia University Press, 2006).
15. See Alan B. Krueger, "To Improve Terrorism Data, the U.S. Should Follow the Lead of Economic Statistics," *New York Times*, July 22, 2004; and Alan B. Krueger and David D. Laitin, "'Misunderestimating' Terrorism," *Foreign Affairs* 83 (September/October 2004): 5–8.
16. Duke University Libraries, "ITERATE—International Terrorism: Attributes of Terrorist Events," June 4, 2009, <http://library.duke.edu/data/collections/iterate.html>.

9

Making Sense of Numbers

Much of a busy executive's workday involves making decisions. These days executives find themselves increasingly surrounded by information, often expressed in numbers. The spectacular advances in technology have made collecting and storing information inexpensive and data retrieval and the display of their analyses quite simple. A large series of numbers, however, tends to numb our senses and push us beyond our cognitive capabilities. Therefore, decision makers prefer seeing large sets of numbers in an understandable form.

The purpose of this chapter is to provide students and practicing analysts with a guide for rendering social, political, or economic phenomena in graphs or tables. It is important to be creative without being misleading or deceptive. I will take you through some familiar terrain and expose you to the advantages and pitfalls of the most commonly used and abused methods of presenting numerical arguments to decision makers.

Descriptive statistics and graphical techniques are useful in assessing social conditions, such as per capita income, the rate of population growth, and crime rates. These techniques are often neglected in more sophisticated statistics and operations research textbooks because they seem too simplistic. However, an overwhelming number of decisions in both public and private organizations are made on the basis of simple decision rules: a brief assessment of relative desirability based on quick impressions rather than thorough research.

A PICTURE'S WORTH: GRAPHICAL METHODS OF ANALYSIS

Let us consider a hypothetical situation. Suppose the city manager of a medium-sized city, Masters, Pennsylvania, would like to know how much the city depends on state and federal grants. The financial management division of the city gives the city manager breakdowns of the city's state and federal grants for 2000 through 2009 (see Table 9.1).

TABLE 9.1**State and Federal Grants to Masters, Pennsylvania, in Current Dollars, 2000–2009**

Year	State and federal grants in current \$
2000	71,000
2001	75,000
2002	85,000
2003	89,000
2004	91,000
2005	91,500
2006	93,000
2007	95,000
2008	98,000
2009	102,000

Current versus Constant Dollars

The numbers in Table 9.1 indicate that state and federal funding for the city is growing every year. A better appreciation of the historical trend can be obtained by plotting the data. However, as you will soon discover, there is more than one way of drawing a picture.

Of course, the plotted information shown in Figure 9.1 is a great improvement over the table. The diagram clearly shows the generosity of the state and federal governments to the city. The dollar amounts of the grants have increased steadily over the years, with the greatest increases coming during the early 2000s. However, the city manager is skeptical; this diagram does not take into account the rate of inflation for the period. Therefore, the price deflator (the Consumer Price Index for the United States) was obtained for the period, and the yearly numbers were converted to constant dollars (see Table 9.2).

Time-series data require the conversion of current dollars into constant dollars because the value of money does not remain the same over time. The books that you purchased last year probably cost more today. Such increases make the value of money

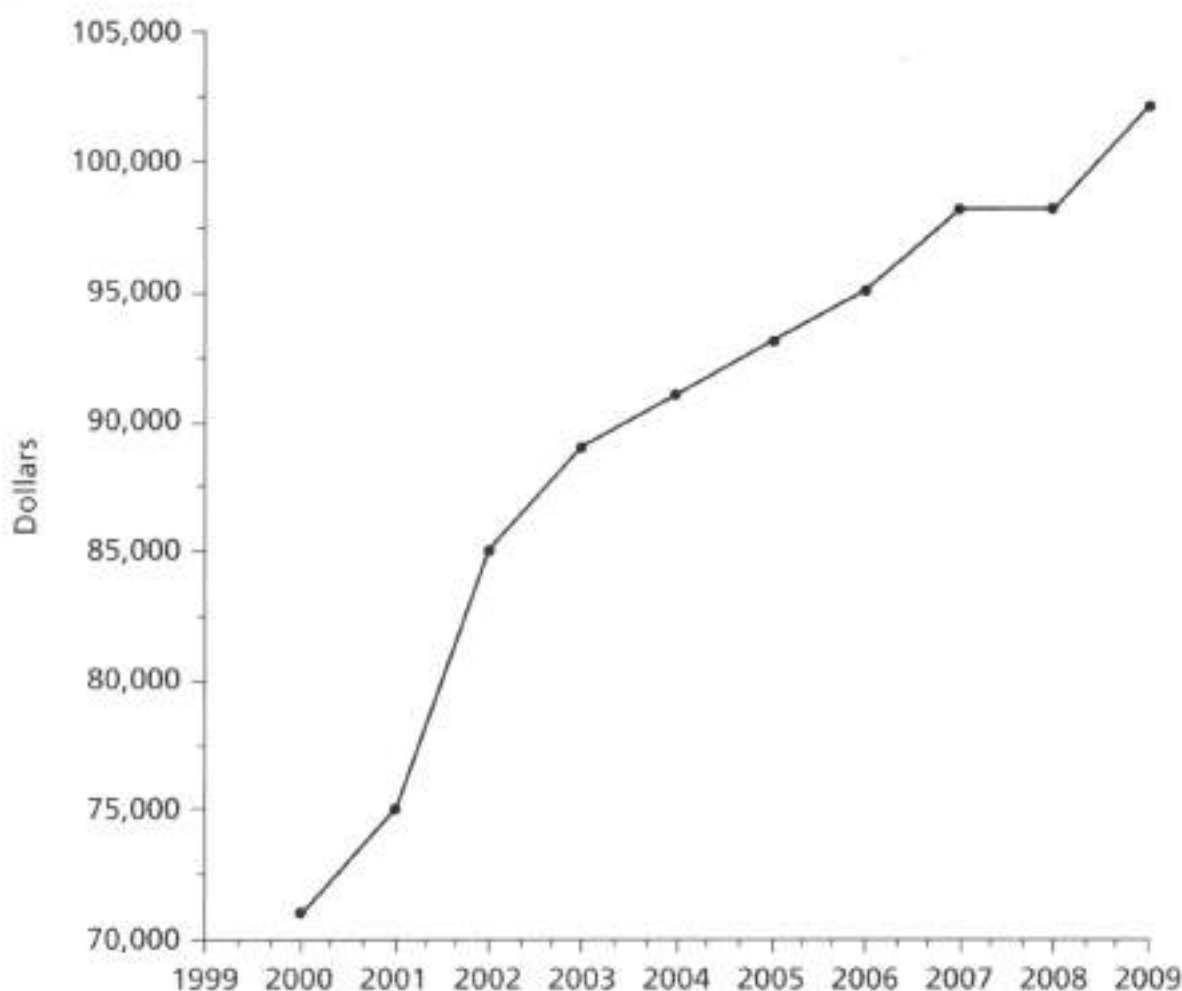
TABLE 9.2**State and Federal Grants to Masters, Pennsylvania, in Current and Constant Dollars, 2000–2009**

Year	State and federal grants in current \$	Consumer Price Index ^a	State and federal grants in constant \$
2000	71,000	100	71,000
2001	75,000	110	68,182
2002	85,000	117	72,650
2003	89,000	122	72,951
2004	91,000	127	71,654
2005	91,500	129	70,930
2006	93,000	137	67,883
2007	95,000	139	68,345
2008	98,000	143	68,531
2009	102,000	147	69,388

^a The Consumer Price Index data are hypothetical.

FIGURE 9.1

State and Federal Grants to Masters, Pennsylvania, in Current Dollars, 2000–2009



decrease with the passage of time. Therefore, in using the actual dollars received by Masters from the state and federal governments, we paint a deceptive picture. The rate at which the dollar loses its value is measured by a value known as a price index. In the United States, the Bureau of Labor Statistics (BLS) collects data on price changes and other employment-related information.¹ Price changes are measured in terms of a base year, which is expressed as a value of 100. For example, most of the current series in price indexes measured by the BLS hold the period from 1982 to 1984 as 100.

A price index is calculated on the basis of what is known as **variable weight** or **fixed weight**. The variable-weight price index states price change as a ratio of a set of goods and services in the current period and their costs in the base year. More precisely, it is written as

$$\text{Variable-weight price index} = \frac{\text{Value of goods at current-year prices}}{\text{Value of those goods at base-year prices}}$$

The effects of price changes are, of course, not universal. If you do not like turnips, you are not affected if their prices go through the roof. However, if you practically live on turnips, your well-being will be affected. If you are making the same amount of money

TABLE 9.3

Consumer Price Index (CPI): All Urban Consumers, 1990–2008

Year	CPI (1982–1984 = 100)	CPI (1990 as the base year)	Annual rate of inflation	Cumulative inflation rate since 1990
1990	130.7	100.0	—	—
1991	136.2	104.2	4.2	4.2
1992	140.3	107.3	3.1	7.3
1993	144.5	110.6	3.3	10.6
1994	148.4	113.4	2.8	13.4
1995	152.4	116.6	3.2	16.6
1996	156.9	120.1	3.5	20.1
1997	160.5	122.8	2.7	22.8
1998	163.0	124.7	1.9	24.7
1999	166.6	127.5	2.8	27.5
2000	172.2	131.8	4.3	31.8
2001	177.1	135.5	3.7	35.5
2002	179.9	137.6	2.1	37.6
2003	184.0	140.8	3.2	40.8
2004	188.9	144.5	3.7	44.5
2005	195.3	149.4	4.9	49.4
2006	201.6	154.3	4.9	54.3
2007	207.3	158.6	4.3	58.6
2008	215.3	164.7	6.1	64.7

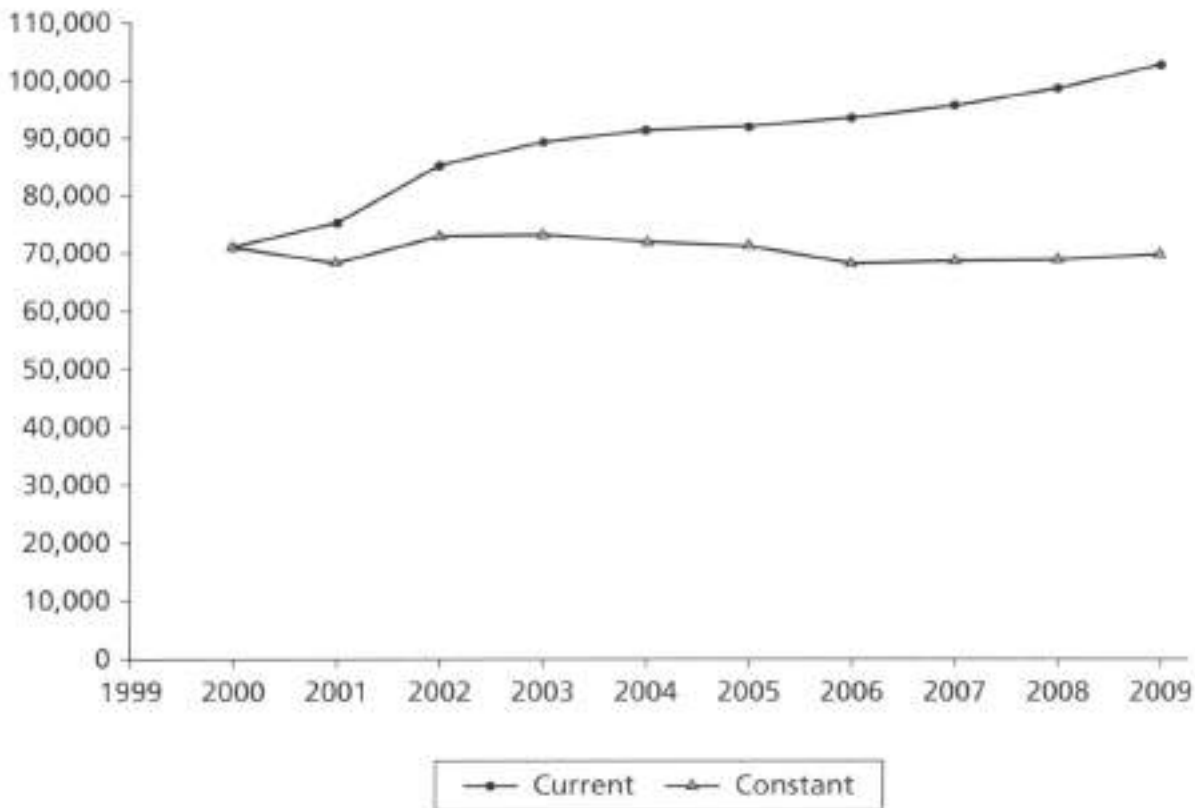
as last year, you are poorer this year than you were before. Your income in current dollars is called your **nominal income**, and the true value of your reduced income is called your **real income** (meaning that it is adjusted for price change). You may note that the variable-weight price index takes an overall picture and does not consider whether the price changes will affect any single segment of the population. As a result, the most commonly used variable-weight price index is called the **GDP deflator**, which is used to compare the gross domestic product (GDP) of the past year with that of the current year.

In contrast, the fixed-weight price index considers a typical basket of goods that a consumer would consume and tracks its prices over the years. The fixed-weight price index is used to measure the **Consumer Price Index**, or the changes in prices that will affect a typical consumer. The BLS produces data series on the Consumer Price Index for various regional centers (urban, rural, a specific city, and so forth) as well as GDP deflators. For example, suppose you have job offers from two different cities. By comparing the price levels of the two, you can decide which one is offering you a higher salary in “real” dollars. In Table 9.3, I show the change in prices facing all urban consumers. From this table you can see that between 1990 and 2008, a dollar lost 64.7 percent of its value. That is, a dollar in 2008 was worth about 35 cents in 1990.

You can also calculate the amount of inflation by changing the base. For example, if you want to use 1990 as the base year, you can recalculate the series using the following formula:

$$\pi_{t+1} = \frac{P_{t+1} - P_t}{P_t} \times 100 \quad (9.1)$$

FIGURE 9.2 State and Federal Grants to Masters, Pennsylvania, in Current and Constant Dollars, 2000–2009



where π_{t+1} is the index of price change for year $t+1$, P_t is the Consumer Price Index for year t , and P_{t+1} is the Consumer Price Index for year $t+1$. The results are shown in Table 9.3 under the heading “Annual rate of inflation.”

Looking again at the case of Masters, it is obvious from the calculation of constant dollar figures that the city is not doing as well with state and federal grants as was assumed. In fact, the inflationary forces in the early 2000s eroded so much of the purchasing power that they caused an actual decline in the grant money in **real terms** (see Figure 9.2).

As you can clearly see, we can get radically different conclusions each time we transform the data. You can further transform the data by comparing them with other variables, such as population or the size of the city budget; the data may also be compared with national or regional averages, or the series can be looked at by its rate of increase. In each case, the data will tell us a different story. In a sense, the data are like a kaleidoscope, in which you can see a completely different picture by slightly changing the angle of the device. Let us consider some other ways of looking at the same information.

We can expand the data presented in Table 9.2 to include information on total city revenue during the period of study. If we express the dollar amounts of grants received by the city as percentages of its total revenue, we will be comforted by the fact that we have done well over the years (see Table 9.4). We can be further comforted if we look at the national trend of the ratio of state and federal government assistance to cities as percentages of their total revenue. While external assistance to cities across the nation

TABLE 9.4**State and Federal Grants to Masters, Pennsylvania, as a Percentage of Revenue, 2000–2009**

Year	State and federal grants in current \$	Total government revenue in current \$	State and federal grants as a ratio of total government revenue	National average of grants as a percentage of local government revenue
2000	71,000	360,000	19.7	35.5
2001	75,000	369,000	20.3	32.3
2002	85,000	382,000	22.3	30.2
2003	89,000	385,000	23.1	26.8
2004	91,000	398,000	22.9	22.8
2005	91,500	410,000	22.3	21.5
2006	93,000	419,000	22.2	20.2
2007	95,000	425,000	22.4	20.3
2008	98,000	432,000	22.7	19.8
2009	102,000	444,000	23.0	19.5

was going down in the 2000s, our city held its ground and was doing better than the national average during the latter part of the decade (see Figure 9.3).

Each of these transformations tells a slightly different story. Through them we get glimpses of different facets of the situation. Therefore, the question is not which one is telling the “true” story but which one contains the most important message from the perspective of the inquirer.

Percentage Change

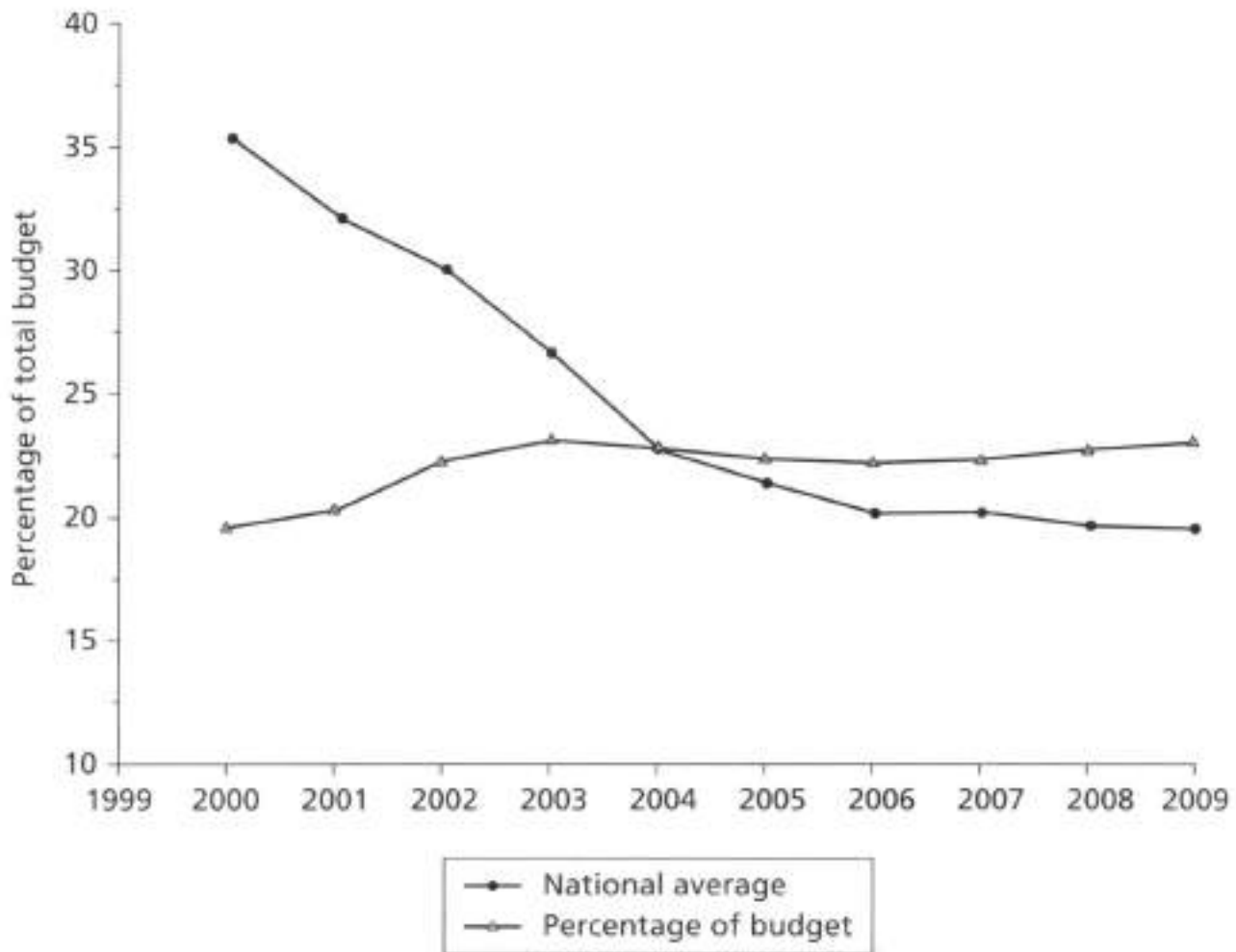
You may want to look at the information in yet another way. You may calculate the yearly percentage change in constant dollar grants to Masters. This information, presented in Table 9.5 and plotted in Figure 9.4, can be quite useful in discerning year-to-year changes in state and federal assistance to the city.

TABLE 9.5**Yearly Percentage Change in State and Federal Grants to Masters, Pennsylvania, in Constant Dollars, 2000–2009**

Year	State and federal grants in constant \$	Yearly percentage change
2000	71,000	—
2001	68,182	−3.97
2002	72,650	6.55
2003	72,951	0.41
2004	71,654	−1.78
2005	70,930	−1.01
2006	67,883	−4.30
2007	68,345	0.68
2008	68,531	0.27
2009	69,388	1.25

FIGURE 9.3

State and Federal Grants to Masters, Pennsylvania, as a Percentage of Local Government Tax Revenue, 2000–2009



Creating an Index

Finally, it may be useful to look at the data with the help of an index. As noted earlier, an index is created when we take a particular figure as the **base** and then express the series in relation to this particular number. For example, if we take the grants figure for 1990 as the base (expressed as 100), we can calculate the index by dividing each year's data by this number and then multiplying it by 100. That is,

$$\text{Index for 1991} = \frac{1991}{1990} \times 100$$

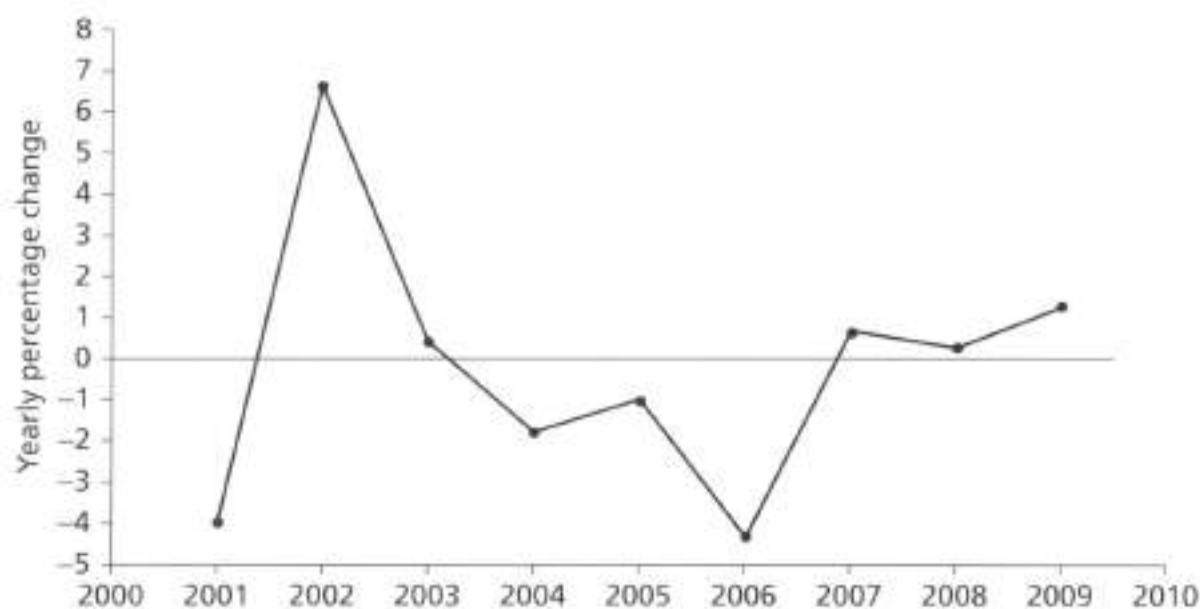
or

$$\frac{68,182}{71,000} \times 100 = 96.03$$

The data presented in Table 9.6 are plotted in Figure 9.5. As you can see, each presentation of the same information tells a slightly different story. Therefore, how you present your case will depend on your need.

FIGURE 9.4

Yearly Percentage Change in State and Federal Grants to Masters, Pennsylvania, in Constant Dollars, 2000–2009



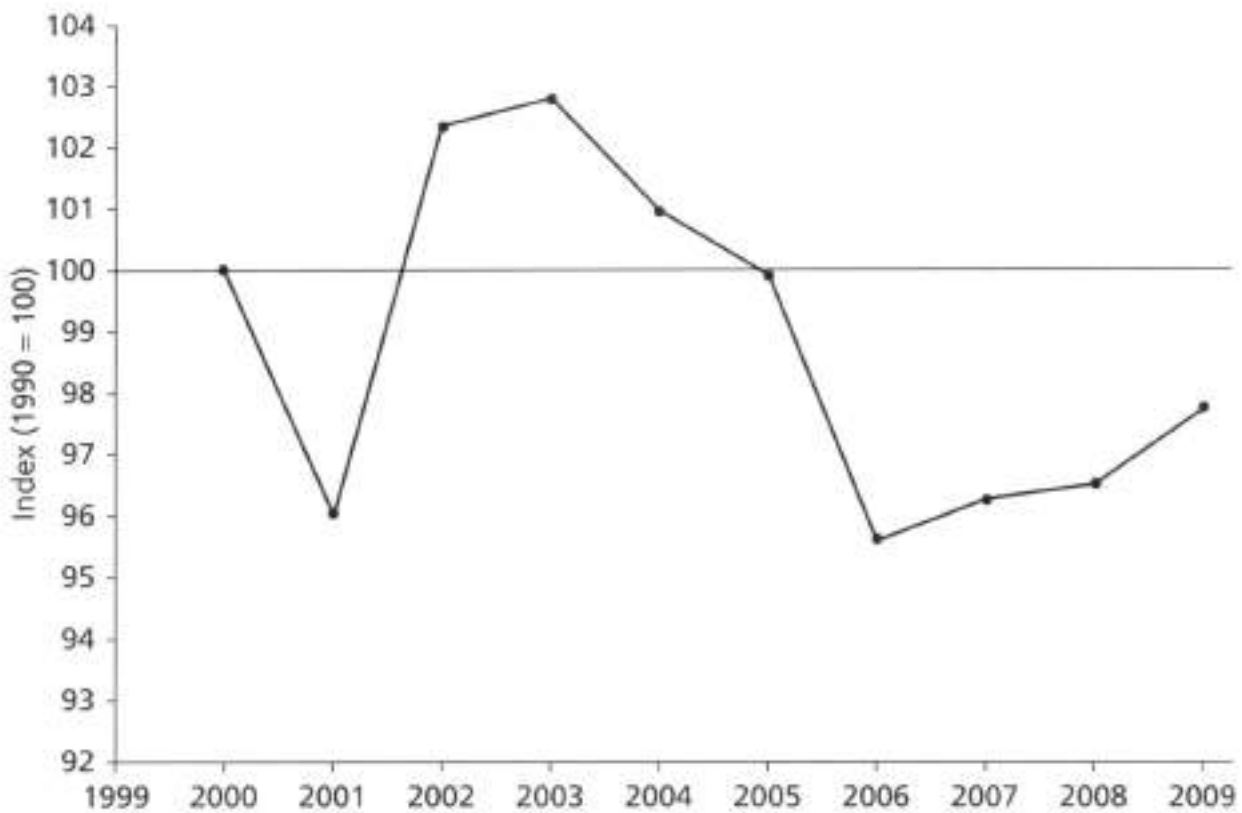
Choosing the Type of Graph to Use

In the previous examples we used only line graphs. Today's managers use a variety of graphs, such as **scatter plots**, **bar graphs**, and **pie charts**. Each kind of graph presents information in its own unique way. You should be familiar with each kind of graph and determine which kind of pictorial rendition gets your intended message across in the most effective way (see "A Case in Point").

TABLE 9.6

Index of State and Federal Grants to Masters, Pennsylvania, 2000–2009

Year	State and federal grants in constant \$	Index (1990 = 100)
2000	71,000	100.00
2001	68,182	96.03
2002	72,650	102.32
2003	72,951	102.75
2004	71,654	100.92
2005	70,930	99.90
2006	67,883	95.61
2007	68,345	96.26
2008	68,531	96.52
2009	69,388	97.73

FIGURE 9.5**Yearly Percentage Change in State and Federal Grants to Masters, Pennsylvania, by Indexing, 2000–2009**

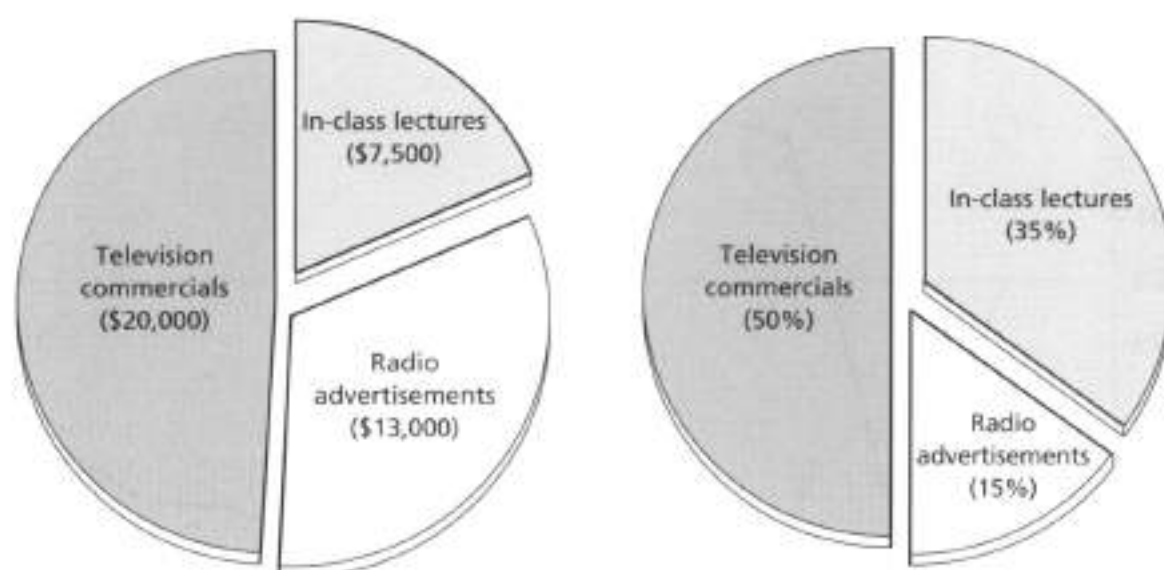
Specifically,

- Line graphs are used for time-series data.
- If you have cross-sectional data, you should use a bar diagram or a scatter plot.
- If the data are in fractions of a total (or in percentages), use a pie chart.

Graphical Methods in Decision Making

Graphical presentations describe a situation by visual means. However, they should not be considered passive tools of description; they can also be used as extremely powerful decision tools. Consider the following situation. The police department in Masters is trying to reach as many youngsters as possible to educate them about the perils of drug use. Last year the department expended considerable effort in arranging school appearances by officers and experts. It also advertised on local radio and television. Suppose that last year, the city spent \$7,500 on school lectures, \$13,000 on radio advertisements, and \$20,000 on local television ads. A recent survey by the city showed that of the children who are aware of the city's drug prevention effort, 35 percent became aware of the issue through face-to-face contact with officers, 15 percent through radio ads, and the remaining 50 percent through watching television. In the two pie charts in Figure 9.6, I show the expenditure for each method of contact.

The side-by-side placement of these two diagrams visually demonstrates that face-to-face contacts are the most cost effective way to disseminate drug prevention

FIGURE 9.6**Expenditures on a Drug Prevention Program in Masters, Pennsylvania**

information to school-age children. The same information could have been presented as simple percentages in a table. However, as is often the case, a picture is worth a thousand words. For another example of how diagrams may enhance presentation of data, see "A Case in Point."

A CASE IN POINT

Racial Profiling

Many minority drivers have protested racial profiling, a practice in which police stop nonwhite motorists and search their vehicles at a disproportionate rate. Mounting concerns about the practice from San Diego's minority population prompted the city's police department to conduct a study. The study covered more than 90,000 instances of police officers' stopping drivers between January and June 2000. The *San Diego Union-Tribune* published the data as pie charts (see charts below).²

The publication of these data created a big stir, particularly when they were presented visually in pie charts similar to the ones here. "We have known this for a long time," said an African American community leader regarding racial profiling. However, in support of police practices, San Diego police chief David Bejarano pointed out that although Latinos made up only about 20 percent of the city's drivers, because San Diego is a border town, a large number of motorists come from Mexico. Furthermore, he pointed out, the figures might be misleading, because police are often deployed in areas close to the border and other high-crime areas. Therefore, considering the entire city population does not allow for a valid comparison. A more detailed study would be required to control for the influence of these two external factors.

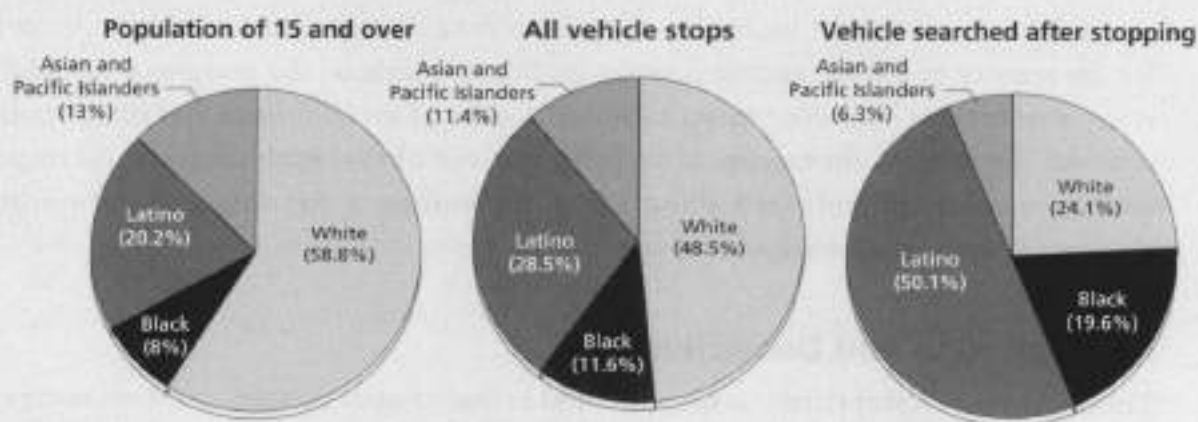
TO TELL THE TRUTH AND NOTHING BUT THE TRUTH

Over the years, statistics have been characterized in less than flattering ways. One of British prime minister Benjamin Disraeli's famous quips was "There are three kinds of lies: lies, damned lies, and statistics." We have come to accept expressions such as "statistical artifacts" and "cooked-up statistics." We must recognize that deception, misunderstood implication, or the existence of a bias in the process of collecting information can cause problems.

The deceptive use of numbers must be defined with respect to the intent of the user. Thus, an individual or organization that puts out information knowing full well that the data have no real-life validity is defrauding or deceiving the user. During times of national emergencies or war, government agencies routinely use data for propaganda purposes. A nation at war may exaggerate or downplay claims about its military or industrial strength, or its war casualties, depending on its strategy. China, in the course of suppressing the prodemocracy movement, underreported the number of student casualties and released figures that were widely disputed by such knowledgeable people and agencies as Chinese student groups in the United States and Amnesty International. Similarly, the figures for North Vietnamese war casualties were routinely inflated by Pentagon officials during the Vietnam War.

Another source of contention frequently centers on the "true" implication of a statistic. We often use per capita GDP as a measure of the relative economic development of nations. However, it is obvious that the word *development* used in a national context should imply more than a measure of a country's per capita GDP, because *development* implies a certain degree of progress and maturity in social, political, and economic

Comparison of Police Stopping, by Race



Discussion Points

1. How effective was the presentation of the data?
2. Can you think of a better way to present the information?
3. What are the points raised by Chief Bejarano? How would you respond to these points?

institutions. Tiny oil-rich nations may have the highest per capita GDPs, but one would be hard-pressed to characterize those countries as the most developed in the world.

Biases resulting from other factors may also cause a statistic to be misinterpreted. Valid questions have been raised about whether IQ tests measure relative levels of intelligence in children. For years, entrance to the U.S. Civil Service was based on one's score on a multiple-choice examination. However, it was eventually determined that such an examination was biased in favor of white, middle-class men. Therefore, the test score could not be accepted as the best measure of a candidate's suitability for a position.

Then there are data that, by their very nature, call for subjective judgment in how they are defined and compiled. A good example is the Consumer Price Index, discussed earlier in this chapter. As mentioned earlier, a "typical" basket of goods and services that an "average" American consumes yearly is used to measure the rate of inflation. However, we know that each of us has a unique consumption pattern, based not only on our individual tastes but also on various factors, such as age, income, race, and geographical location. If the price of skateboards goes up, senior citizens are less likely to be affected than are young people. Similarly, an increase in the cost of health care may not affect single young adults as much as the increase in the index would suggest. Whoever compiles this basket of goods thus faces two problems in making the index relevant to the majority of Americans. First, the compiler must discern what is "typical," in terms of what kinds of goods and services and at what levels of consumption. Errors that severely distort the data creep in during the collection of information in many ways; not recognizing this means that the results of analyses may be meaningless, misleading, or even damaging.

You must be extremely wary of accepting data for analysis, and you cannot be too careful in looking at possible sources of bias and errors. At the same time, remember that it is impossible to find a perfect set of data in an otherwise imperfect world. Like the proverbial fastidious eater who dies of starvation, a researcher who is too cautious will know all the flaws of the data and its analyses without being able to draw any useful conclusions from them.

The rule of thumb, then, is to evaluate carefully the sources of bias in the data and be aware of the cost of doing an incorrect analysis. If you are conducting medical research for a new type of vaccine as an antidote for a disease, or calculating trajectories for the reentry of a space shuttle into the earth's atmosphere, the margin of acceptable error is rather low. However, mercifully, in the areas of social science and public policy research, the demand for numerical accuracy may not be that extreme. It is most important to be aware of and open about the shortcomings of the data and the possible sources of bias in their analysis and interpretation.

Interpretation and Deception

The last source of skepticism to keep in mind is that to most people, numbers portray a rigid, self-evident truth. In a cocktail party discussion, a friend claimed that homosexuality was purely biological, since homosexuals seem to constitute 10 percent of every society's population. It is fairly obvious that this statement is the kind designed to end all discussions, as it purports to present a totally scientific, incontrovertible fact of life. To some people, numbers pose a threat because they have the appearance of "scientific" objectivity. However, closer scrutiny will reveal problems resulting from the various biases we have described, and a significant source of disagreement may be that any

information (numerical or otherwise) about a complex social situation is bound to be open to interpretation.

In 1954 Darrell Huff wrote an extremely interesting, humorous book, *How to Lie with Statistics*, in which he systematically demonstrated many ways to distort information to suit the purposes of an investigator. In his tongue-in-cheek introduction he states,

This book is a sort of primer in ways to use statistics to deceive. It may seem altogether too much like a manual for swindlers. Perhaps I can justify it in the manner of the retired burglar whose published reminiscences amounted to a graduate course in how to pick a lock and muffle a footfall: The crooks already know these tricks; honest men must learn them in self-defense.³

Huff's highly acclaimed book advanced understanding of the various ways one can use descriptive statistics among generations of undergraduate students. However, in all honesty, we may pose the question differently. If the manipulation of data is always suspected of "distorting" the picture, then there must be a truly undistorted version of real life. In other words, are we to assume the universality of truth? Does it always require a statistician to obfuscate an otherwise obvious situation? A famous early twentieth-century Japanese play *Rashomon*, by Ryunosuke Akutagawa, brings home the point of the relativity of perception. In the play a bandit rapes a young woman traveling with her samurai husband. A number of different individuals witness this terrible act of violence. When they are brought to the trial (including her deceased husband, who speaks through a medium), the incident is found to have variations of interpretation. As the play shows, there may be honest differences of opinion in the way one looks at a situation, even when expressed in "cold, hard, objective numbers." We live in a complex world in which "truth" may have more dimensions than can be effectively captured by any one-dimensional measure. However, if we use multiple indexes to characterize a situation, our cognitive limitations stand in the way of formulating any definitive picture. Like everything else in life, the quantification of social phenomena requires a trade-off between the confusion of a total picture and the clarity a limited view offers.

For example, consider our hypothetical city of Masters, Pennsylvania. The demographic composition of Masters is typical of the region, with a large number of working-class people, along with pockets of urban blight, characterized by persistent levels of high unemployment. However, a few areas of the city house extremely wealthy families. Let us focus on three individuals plying their trades in Masters: a real estate broker, a college professor, and a city planner. All three of these individuals want to present the "true" economic picture of this city with a single number: the average income of the city's people. However, these three have different objectives. The real estate broker wants to portray the city as a nice place to live and raise a family. Therefore, in talking to clients he mentions as "average" the mean income of the residents of the town. However, although small in number, the extremely wealthy households influence the mean. The prospective buyer gets a much rosier picture of the average affluence of the city than that espoused by the professor. The professor is conducting research in urban economics, for which he is using the figure of median income. The median, the middle income across the range from highest to lowest, presents a less attractive picture of the economic well-being of the city because it is not affected by the presence of the wealthy sector of the community. However, even this number is far superior to the one used by the city planner. The city planner of Masters wants to respond to a request for a grant proposal from



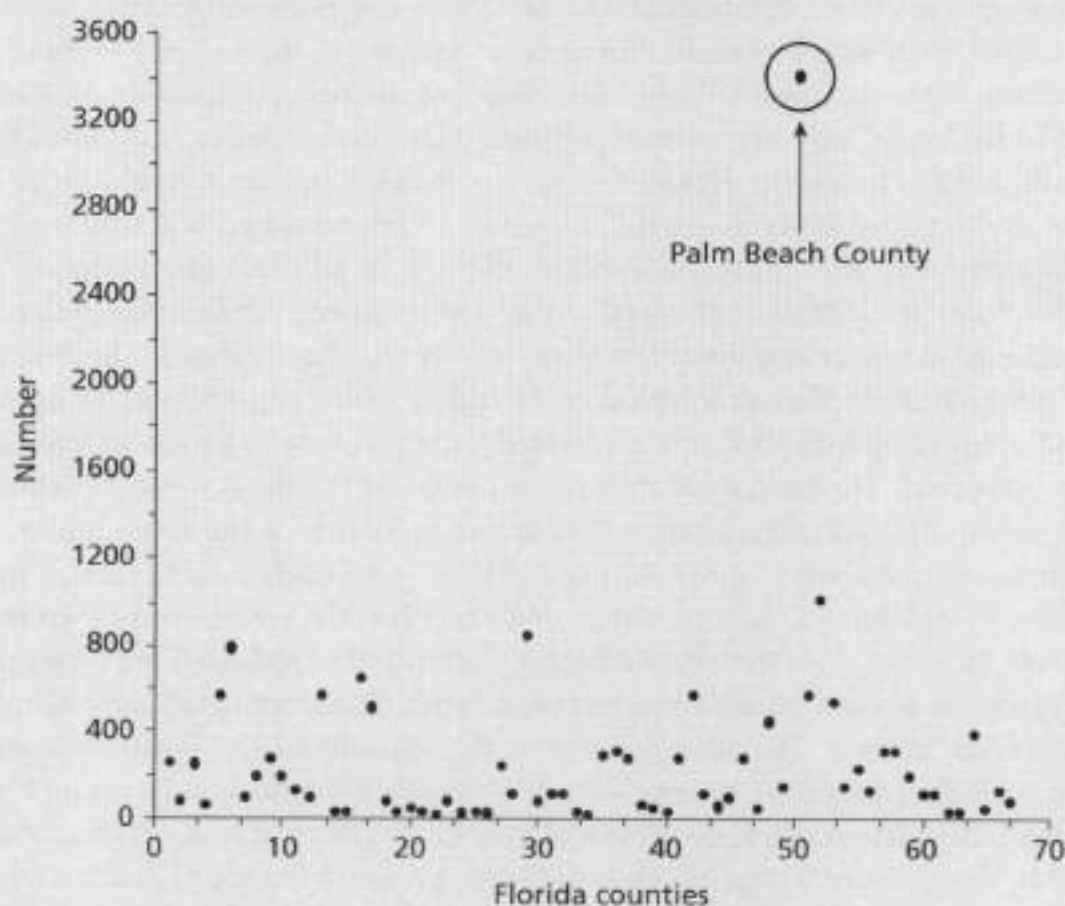
A CASE IN POINT

Digging Deep into the Numbers: A Stolen Election?

The 2000 presidential election looms large in the annals of modern American history. Many accused the Republican candidate, George W. Bush, of stealing the election from the Democrat, Al Gore. Others saw the results differently.

As the votes were tallied, in a neck-and-neck contest, Florida took on added prominence because the state's twenty-five electoral votes would determine the outcome of the election. Within the state of Florida, the voting of Palm Beach County residents generated an unexpected twist. In the 1992 presidential election, Ross Perot, the candidate of the newly founded Reform Party, drew a respectable number of votes, which many thought gave the presidency to Bill Clinton. In 2000 the Reform Party candidate was Patrick Buchanan. In Palm Beach County, the election commission approved a "butterfly" ballot, a double-faced ballot on which the names of the candidates appeared side by side. The name Al Gore was opposite that of Pat Buchanan, creating confusion among many Democratic voters. Many Gore supporters mistakenly voted for Buchanan instead. When the votes were tabulated, Buchanan had received about 3,400 votes.

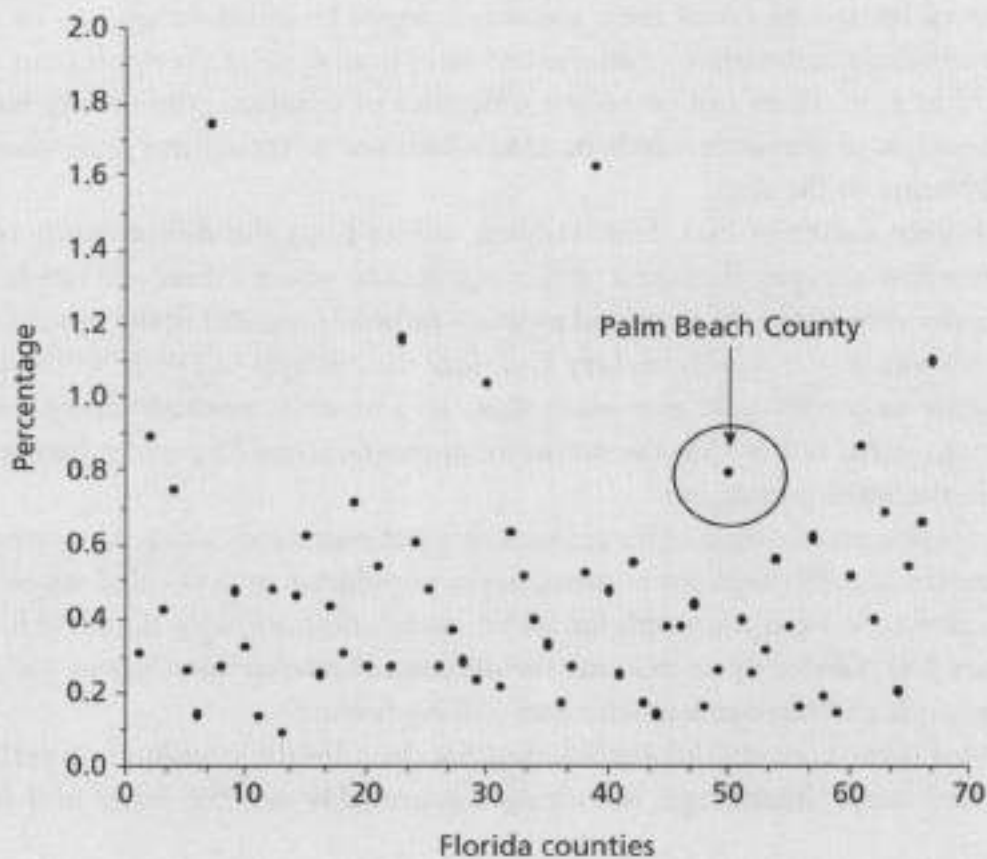
Numbers of Votes Cast for Patrick Buchanan in the Florida Counties



The graph above shows the number of votes garnered by Buchanan in Florida's sixty-seven counties, before they were hand counted, as a scatter diagram. As you can see, the numbers clearly show an unmistakable spike for Mr. Buchanan. The scatter plot clearly identifies those Palm

Beach Buchanan votes as an obvious outlier. However, in the social sciences, for every obvious solution, there is always a different angle. When we calculate the number of votes as a percentage of the total votes in each county, Palm Beach County's record does not appear so strange, as the figure below demonstrates. In fact, in terms of percentages, at least six small counties (Baker, Charlotte, Indian River, Liberty, Suwannee, and Washington) polled higher than did Palm Beach.

Percentages of Votes Cast for Patrick Buchanan in the Florida Counties



Discussion Points

1. The importance of the U.S. election cannot be overemphasized in terms of its impact on the world stage. In 2009, U.S. Secretary of State Hillary Rodham Clinton went to Pakistan and held a town hall-style meetings with university students. When Clinton pointed out the importance of having legitimate, democratically elected governments in Islamabad and Kabul, some students pointed to the controversial Florida election results as an example of imperfect democracy in the United States. Although there are many aspects to the truth, after reading this case, what conclusion can you derive about the legitimacy of the election in Florida?
2. What lesson can we derive from looking at the two scatter plots about having a healthy skepticism about any data set?

the state government to bring in money earmarked for the economically depressed areas. For this proposal, she uses the modal income of the town, which is the most frequently found income of its inhabitants.

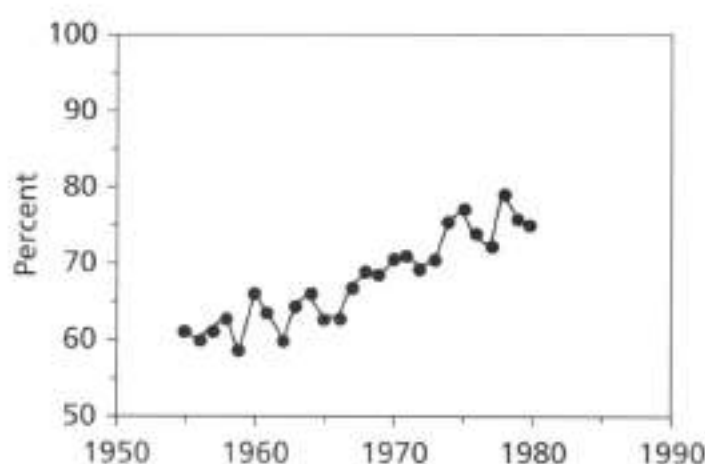
Is it possible to pick out which of the three individuals, who use three different measures of average income because of their different objectives, is engaged in an act of deception? I would argue that none of them can be accused of such an act unless some other kind of deception is present. When it comes to the definition of “average,” most people intuitively use the arithmetic mean, median, and mode, in that order. Therefore, by convention, if one uses the term *average* for the mean, one can feel justified. The use of the median may require justification, and the use of the modal income would certainly require its mention in the report, to be ethically fair and aboveboard. However, the use of any of these measures cannot be called deception. Therefore, we must conclude that without the intent of deception, none of the figures can be characterized as a lie. There can be honest difference of opinion, even among those whose business it is to deal with numbers, as to which one of these three represents the most valid picture of the city.

Another source of bias, Huff claimed, comes from the deceptive use of pictorial information—graphs. Because a picture is said to be worth a thousand words, the desire to convey information by graphical means is rather strong. But in the process one might take advantage of a certain trickery. Consider the example in Figure 9.7, which depicts nonwhite unemployment as a percentage ratio of white unemployment. In 1955 the unemployment rate within the nonwhite population was 62 percent higher than that within the white population.

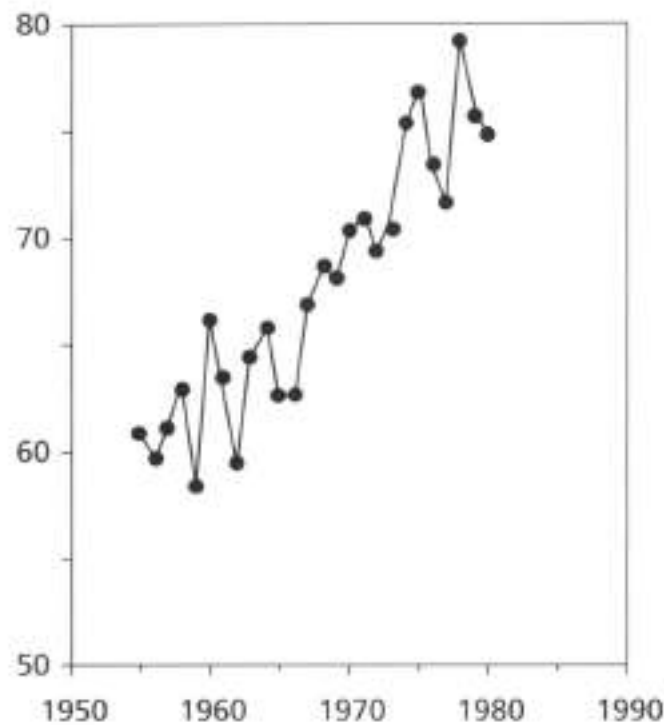
Does this presentation of the information suit your needs, or do you want to portray more dramatically the plight of the minority population in the United States? If you do, you can *s-t-r-e-t-c-h* the graph for added visual effect, showing the same information (Figure 9.8). Obviously, in this case the difference between nonwhite and white unemployment is portrayed in a much more striking fashion.

What if you contend that the situation for the minority population is really not that bad, or that the situation has not changed appreciably over the years? In that case, you

FIGURE 9.7 Ratio of Nonwhite to White Unemployment Rate



Source: U.S. Bureau of Labor Statistics.

FIGURE 9.8**Ratio of Nonwhite to White Unemployment Rate
(stretched graph)**

can use another trick. You can increase the range of the vertical axis, which allows you to present the same information in a different light. Against a much wider range of possible ratios of unemployment, the ethnic difference in the relative measure of economic deprivation does indeed look small (Figure 9.9).

Yet another effective way of representing a series is to use selective years for comparison. For example, you may want to present the same information contained in the three graphs, but even more emphatically. You may want to show them in either of the following ways without “lying” with your statistics. The two graphs in Figure 9.10 show only three years: 1960, 1970, and 1980. This restriction removes the distracting effects

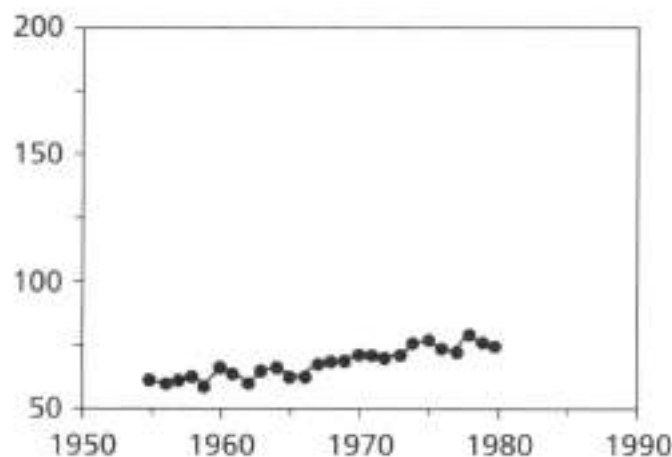
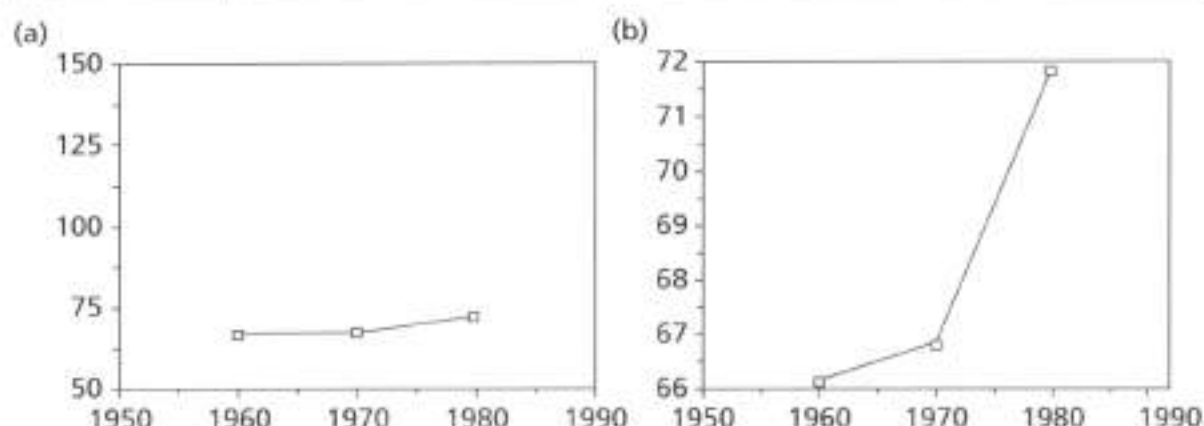
FIGURE 9.9**Ratio of Nonwhite to White Unemployment Rate
(elongated vertical axis)**

FIGURE 9.10**Ratio of Nonwhite to White Unemployment Rate
(1960, 1970, and 1980)**

of yearly fluctuations and allows us to present long-term trends. Then, by simply manipulating the vertical axis, we have two radically different visual effects.

Now that I have shown you various ways of presenting the same information, which do you think represents the “true” picture? The answer is simple: we do not know which of these diagrams would be classified as a deceptive representation of the reality. However, Figures 9.9 and 9.10 (a) might be interpreted as edging toward questionable practices, since the vertical axes in those figures have wider ranges than required by the data. But would you call that lying? In real life, truth, like beauty, is in the eye of the beholder.

In the preceding examples, the difference between interpreting and deceiving might have been subtle, but consider the rendition of the same information in Figure 9.11.

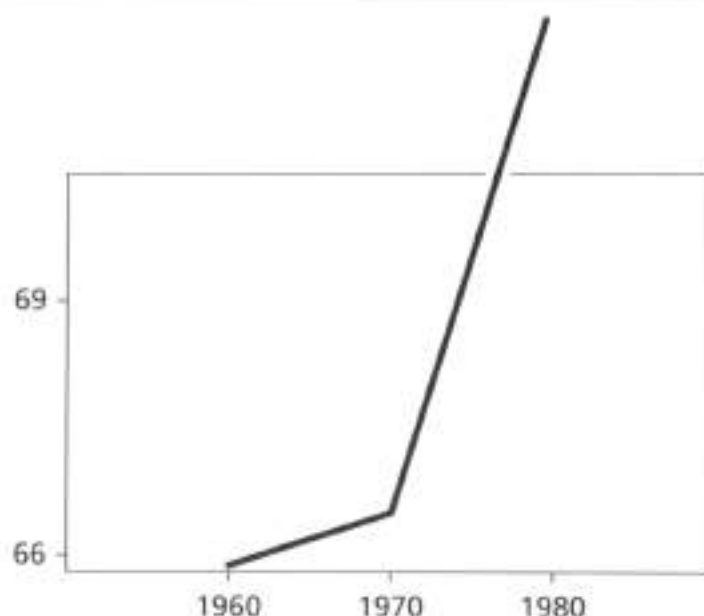
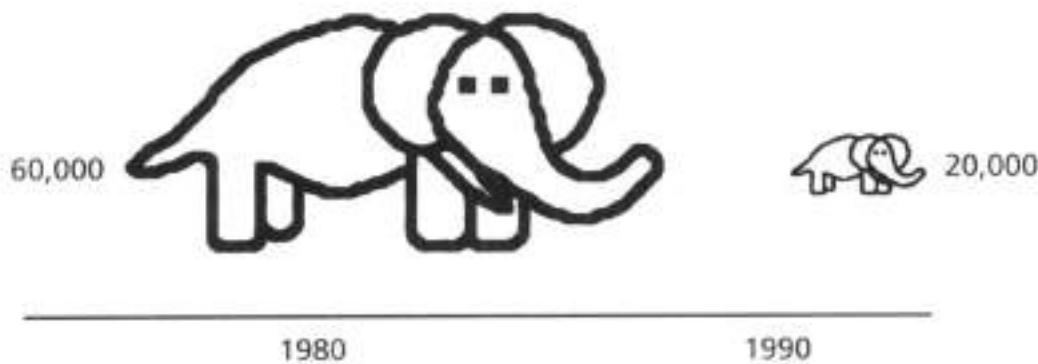
FIGURE 9.11**Ratio of Nonwhite to White Unemployment Rate
(1960, 1970, and 1980; shortened vertical axis)**

FIGURE 9.12 African Elephant Population, 1980–1990

In this diagram, you are not only presenting facts but also trying to make a rather loud statement with the obvious analogy of the ratio going through the roof.

Huff pointed out that, intended or unintended, deception may creep in more readily when one uses graphics instead of lines and bars, which are drawn according to scale. Consider a diagram showing the plight of African elephants (Figure 9.12). It is shocking to realize that in ten years the elephant population has diminished to one-third of its 1980 size. Although the numbers themselves speak volumes about the plight of the hapless pachyderms in Africa, the pictures in the figure are not drawn according to scale. We catch the relative size difference in the picture more readily than the difference in magnitude between the numbers, but we may not be able to catch that the 1980 drawing is much more than three times the size of the 1990 drawing. These kinds of disproportionate pictorial renditions are popular with those who present data at legislative hearings and in administrative decision-making sessions because the message is direct and dramatic. With computer graphics so easy to use, the possibility of some creative minds' deceiving the unwary has never been greater.

The power of strong visual presentation is well recognized and well documented. A 1985 *Wall Street Journal* story documents how the use of such a tool helped Caspar Weinberger, then secretary of defense, avert a cut in the Pentagon budget by David Stockman, director of the Office of Management and Budget:

In staving off Mr. Stockman's assault on the planned buildup, Mr. Weinberger turned to a tactic for which he has since become famous, the chart and easel. The defense secretary's charts, presented in a meeting with President Ronald Reagan, showed large soldiers bearing large weapons, which were labeled "Reagan budget." They towered above small soldiers with small weapons labeled "OMB budget." President Reagan went along with the "Reagan budget."⁴

Tabular Presentation of Data

A long series of numbers pushes us to the limits of our cognitive capacities. For that reason, the value of an effective table as a means of communication cannot be overstated. Presenting data in useful tabular form is an art, which can be perfected only through practice.

TABLE 9.7 Spending by Function and Level of Government, 1995

Function	Spending by level of government (in million \$)			PS (percentage)	Percentage of spending by level of government			
	Federal	State	Local		Federal	State	Local	Total
Defense	327,231	—	—	9.60	100.00	—	—	100.00
Education	27,270	101,510	276,763	12.10	6.72	25.03	68.25	100.00
Highways	731	48,893	30,216	2.38	0.92	61.24	37.85	100.00
Welfare	57,246	160,421	32,669	7.47	22.87	64.08	13.05	100.00
Police	7,563	5,735	52,329	1.96	11.52	8.74	79.74	100.00
Health	26,517	49,487	56,460	3.95	20.02	37.36	42.62	100.00
Administration	19,416	24,781	35,237	2.37	24.44	31.20	44.36	100.00
Insurance	558,291	93,692	13,648	19.86	83.87	14.08	2.05	100.00
Other	680,857	402,563	262,046	39.93	50.60	29.92	19.48	100.00
Total	1,705,122	887,082	759,368	100.00	50.88	26.47	22.66	100.00

Sources: U.S. Department of Commerce, Census Bureau, and The Tax Foundation (Washington, D.C.).

Notes: PS = percentage of total public spending. Some percentages may not add up to 100 percent because of rounding.

The preparation of an effective table requires considerable thought and time (and therefore money). First, have a clear idea of exactly what you want to communicate to the reader. Second, choose a title that succinctly describes the contents of the table. Third, consider various ways of presenting the raw data so that they make the point you want to make most effectively.

For a table to be useful, its purpose must be absolutely clear from its title. The title should be concise but not so brief that it does not convey the true intent of your presentation. The numbers, taken together, must tell a coherent story. Consider, for example, Table 9.7. You may notice how much information has been packed into a concise table. You can find out from the table how much money the government spent on selected items in 1995. During that fiscal year, education accounted for 12.10 percent of all public spending, and local governments paid the bulk (68.25 percent) of education costs. The table also tells you that the federal government originated half (50.88 percent) of total government spending.

In my discussion of graphical presentation, I have shown you how to look at a data set from different perspectives. Similarly, before preparing a table, you may draw interesting conclusions by using absolute numbers, ratios, percentages, and so on.

THOSE NOT-SO-INNOCENT NUMBERS

It is relatively easy to define outright deception or lying by the measure of intent and the sheer fabrication of data, but the line between deception and differences in interpretation is murky. Often the intentions of presenters are not obvious; nor are we capable of detecting purposeful contamination of data. Because we tend to believe in the objectivity of numerical information more readily than in the subjectivity of qualitative statements, the deceptive use of statistics can bring incredible misery to people.

In 1896 Frederick L. Hoffman, a nationally famous statistician for the Prudential Insurance Company of America, wrote a book titled *Race Traits and Tendencies of the American Negro*. Hoffman's thesis was that since emancipation, African Americans (having left the protective care of their slave owners) had gone back to their "basic racial trait" of "immorality of character." Hoffman based his theory on a number of different statistics he had collected. He noted that in 1890, there were 567 blacks in prison for rape, which constituted 47 percent of the prison population convicted on rape charges. Because this number was significantly greater than the proportion of the African American population (about 10 percent at the time), according to Hoffman, rape and other sexual crimes were reflective of the "Negro racial trait." Hoffman thus concluded that

all the facts brought together in this work prove that the colored population is gradually parting with the virtues and the moderate degree of economic efficiency developed under the regime of slavery. All the facts prove that a low standard of sexual morality is the main and underlying cause of the low and anti-social condition of the race at the present time.³

Hoffman then connected the "Negro racial trait of immorality" to the high mortality rate among the black population. On the basis of this causal linkage, disregarding the fact that the census of 1890 showed a steady increase in the size of the black population, Hoffman predicted that African Americans were doomed to face a "gradual extinction of the race." The name of Hoffman's publisher, the American Economic Association, added a dose of respectability to this statistical study, which was widely used as a weapon in promoting white supremacy for decades to come. However, another important consequence of this and other internal statistical studies was that Prudential judged African Americans to be bad actuarial risks and promptly started to cancel all African American life insurance policies. Within four years, by the end of the century, most insurance companies had gotten out of the business of insuring African Americans.⁶ In a similar manner, statistics have been used over the years to perpetrate many kinds of heinous crimes, or their faulty uses have led to extremely inefficient public policies.⁷

KEY WORDS

Bar graphs (p. 210)

Base (p. 209)

Consumer Price Index (p. 206)

Fixed weight (p. 205)

GDP deflator (p. 206)

Nominal income (p. 206)

Pie charts (p. 210)

Real income (p. 206)

Real terms (p. 207)

Scatter plots (p. 210)

Variable weight (p. 205)

EXERCISES

1. Write an essay on truth and objectivity in quantitative analysis for public policy. Within this context, describe the relative advantages and disadvantages of the various measures of central tendency and dispersion. Provide appropriate examples.
2. After reviewing the U.S. Government Accountability Office report on poverty in Appendix A, consider the yearly poverty rate as measured by the U.S. Census Bureau

from 1959 (see www.census.gov/hhes/www/poverty/histpov/hstpov2.html). Plot the data and examine the trend. Suppose you are an antipoverty advocate; is there any way you can present the information to show that the situation is worsening? On the other hand, suppose you are an ardent proponent of the free market. Can you use the same series and show that the poverty rate is either unchanged or heading downward?

3. Collect data on the growth rate of per capita GDP, the rate of inflation, and unemployment from 1950 (consult the *Economic Report of the President* and the *Statistical Abstract of the United States*). First make the case that the nation has been better served by Republican presidents, and then make the case for the Democrats on the basis of the same set of data.
4. Look at some recent news reports presented with quantitative data. See if you can derive a different conclusion from the same information.

Notes

1. You can access data collected by the BLS at its Web site, <http://stats.bls.gov>.
2. Mark Arner and Joe Hughes, "Police Stop Blacks, Latinos More Often: Data from Profiling Report Echo Fears of S.D. Minorities," *San Diego Union-Tribune*, September 29, 2000.
3. Darrell Huff, *How to Lie with Statistics* (New York: W. W. Norton, 1954), 9.
4. Tim Corrigan, "Weinberger Finds His Well-worn Strategies Always Succeed in Blunting Defense Budget Axe," *Wall Street Journal*, March 1, 1985.
5. Frederick L. Hoffman, *Race Traits and Tendencies of the American Negro* (New York: Macmillan, 1896), quoted in Joel Williamson, *The Crucible of Race: Black/White Relations in the American South Since Emancipation* (New York: Oxford University Press, 1984), 329.
6. For a detailed discussion, see Williamson, *The Crucible of Race*.
7. For an excellent discussion of measuring people's abilities with numbers, see Stephen Jay Gould, *The Mismeasure of Man* (New York: Norton, 1981).

Projection Techniques: When History Is Inadequate

PROJECTION VERSUS CAUSAL PREDICTION

One of the most important of an organization's functions is planning for the future. How would you begin the process? It is not merely a cliché that it is impossible to predict the future without knowledge of the past. In history we look for **trends** and the **causal connections** that offer explanations for the events at hand. We may call the analysis of trends **projection**. In contrast, **prediction** is an inquiry into the causal relationship that binds the variable to be explained (the **dependent variable**) to a set of variables (the **independent variables**) that purport to explain it.

We call trend analysis "projection" because it contains an underlying hypothesis that whatever factor(s) set in motion a past pattern of change will continue to operate in the future, leading to the same rate of growth or the same pattern of behavior. This postulation, called the **assumption of continuity**, is the underlying premise of all forecasting methods. In using projection techniques we aim to find a past trend and project it into the future. Of course, forecasting on the basis of past trends raises the philosophical question of whether history really repeats itself. Without getting embroiled in this age-old controversy, we can safely point out that because progress in the realm of the social sciences is mostly evolutionary and incremental, the study of any long socioeconomic series would point to the existence of some sort of trend pattern.

With causal analysis, in contrast, we hypothesize that the future development of the dependent variable is not related to its past trend (at least not to any significant extent), and therefore, a past pattern cannot solely predict future behavior. Instead, it is determined by a complex causal linkage between the dependent variable and a set of independent variables.

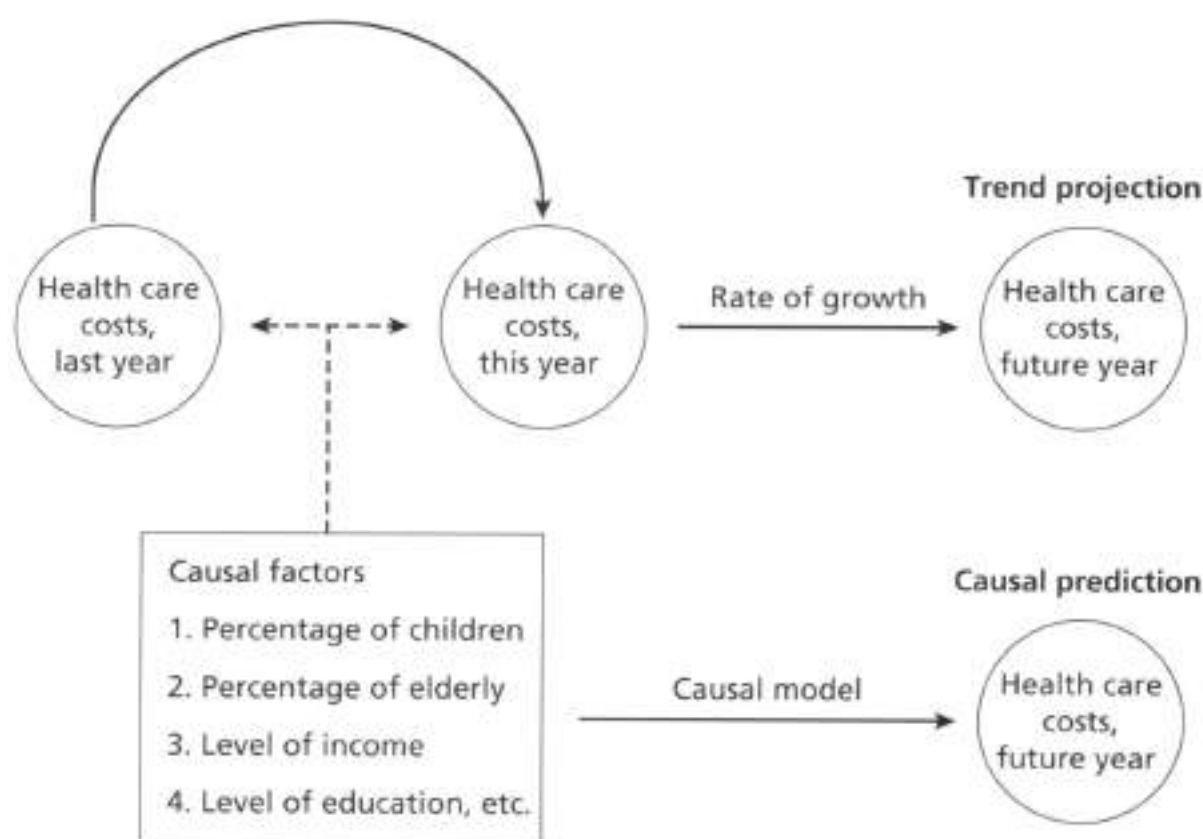
An example may clarify the difference. If we want to know the extent of future health care needs, we can look at past trends and see that health care costs increased by a certain percentage every year and

then forecast that in five years, we will need a corresponding amount of money to meet health care needs. Or we may look at various explanatory factors, such as the percentages of children and elderly in the population, expenditures on preventive medicine, trends related to food and nutrition, education levels, or the rate of income growth. Then, by developing a model in which all these factors influence the outcome, we can attempt to forecast future health care costs (see Figure 10.1).

Both projection and prediction methods have their relative advantages and drawbacks. Projection methods often turn out to be easier to use for forecasting—particularly for short-term “policy prescription”—because these models require data points of a single series to calculate the trend. In contrast, causal models require a thorough understanding of the causal linkages between a dependent variable (outcome) and a set of independent variables (causes). These types of models are better suited for policy research and for longer term forecasting. In general, when choosing an appropriate method of forecasting, you should consider the following points. You should choose a trend analysis in preference to a causal model if

- the past trend is stable, which means not much has changed over time, at least in the short run;
- you do not have a good understanding of the causal factors explaining the dependent variable; or
- you do not have a great deal of time and resources to conduct the appropriate research.

FIGURE 10.1 Trend Projection versus Causal Prediction



INADEQUACY OF HISTORY

If we want to forecast the future with the help of the past, we need a minimum amount of information from the past to provide us with any kind of meaningful insight into the future. Unfortunately, researchers and policy analysts often face the problem of inadequate past information. This inadequacy may result from (1) a lack of history itself, (2) a past lack of interest in collecting information, or (3) scattered past information whose compilation in a series would require too much time and money.

For example, if you are attempting to forecast population growth for a newly incorporated city, you do not have much of a history to fall back on. Or in an attempt to estimate the influx of tourists in a city, you may find out that such data exist for only the past two years or that no systematic effort to collect them has ever been made. Or you may discover that although various agencies have collected the data over a number of years, each has done so independently, without cooperating much with the others; as a result, the compilation of those data in one continuous series would require a great deal of resources. In such cases, you can still make a projection on the basis of one or two past data points, or none at all. For these kinds of cases, I discuss a number of techniques that can be classified under two broad headings: **single-factor projection** and **judgmental methods of projection**.

Single-Factor Projection

Suppose we are interested in projecting the number of cases of AIDS in our city, which was incorporated a couple of years ago. Our town has data on the number of AIDS victims for only the past two years. Last year, there were 150 reported cases, and the year before, there were 138 cases. Therefore, the rate of AIDS in our area increased by nearly 9 percent last year. On the basis of this information, we may project next year's infection rate by multiplying the number of this year's cases by 1.09. In symbolic terms, this can be written as

$$P_1 = (1 + r) \times P_0 \quad (10.1)$$

where P_1 is the future year's population, r is the rate of population growth, and P_0 is the present year's population.

Using this formula, we can predict that if the current population of AIDS victims is 150, one year from now, this number will reach 150×1.09 , which is equal to about 164. Using the same logic, we can see that the year after, this number will increase by another 9 percent. Thus, if P_2 is the population for the second year, then it can be estimated as

$$P_2 = (1 + r) \times P_1 \quad (10.1a)$$

Since we know from equation 10.1 that P_1 is equal to $(1 + r) \times P_0$ we can substitute this expression for P_1 to obtain

$$P_2 = (1 + r) \times (1 + r) \times P_0$$

which means

$$P_2 = (1 + r)^2 \times P_0 \quad (10.2)$$

Therefore, it is obvious that if the population grows at a constant rate, to estimate the second year's population we must multiply the current year's population by a factor of 1 plus the rate of growth to the power of 2. If the population keeps growing at this constant rate, in three years we can expect the number to reach the current year's population multiplied by a factor of 1 plus the rate of growth to the power of 3. Therefore, by generalizing this logic, we can write the formula as

$$P_n = (1 + r)^n \times P_0 \quad (10.3)$$

where n is any year in the future.

Thus, if we want to estimate the number of people who might be infected with the deadly virus ten years from now, we set the number of years equal to ten in the preceding equation ($n = 10$). Thus, we write equation 10.3 as

$$P_{10} = (1 + 0.09)^{10} \times P_0 \quad (10.4)$$

Using this formula, we can estimate that ten years from now, *if the present rate of infection continues*, we can expect to see 355 people infected.¹ This is the formula of geometric growth, which is also used to calculate compound interest rates. The use of this formula can make an analyst's task easy because the size of any future population can be estimated simply by substituting the appropriate numbers in equation 10.3.

For policy analysis, it is often important to know not only the point estimate of the number of cases n years in the future but also the total number of cases for the entire time span. To estimate the total number of people who will develop AIDS during the next ten years, we can estimate the size of the afflicted population for each year using equation 10.5 and then adding these values up. Since this process is cumbersome and time-consuming, we can estimate the sum of any geometric series using the following formula:

$$\sum_{n=1}^k P_n = \frac{(1 + r)^{k+1} - 1}{r} \times P_0 \quad (10.5)$$

Using this formula, we can estimate with a hand calculator the total number of patients needing treatment and public assistance to be nearly 2,634.² By multiplying this number by the current level of medical costs and public funding per patient, we can provide a rough-and-ready estimate of the total medical costs for these potential patients and the amount of money required in public assistance for the next ten years.

In the preceding example, we projected the future values of a variable on the basis of its past rate of growth. However, we can use many other factors as the basis of projections. For example, suppose a large parcel of vacant land is being considered for rezoning. The new zoning ordinance will allow residential housing, apartments, or commercial buildings. Analysts may use past information on land use patterns and their fiscal impact to project the future needs of the new community. By looking at the prices of the

proposed residential units, planners can get a good idea about the economic capabilities of the newcomers. From the data of a larger city or a similar neighborhood, projections can be made for detailed demographic characteristics, such as the number of school-age children and elderly residents, commuting behavior, and recreational needs. If we expect 10,000 people to settle in the new neighborhood, by calculating the percentage for each demographic segment, we can project their numbers. These projections in turn would allow the projection of future resource needs for schools, libraries, recreational facilities, roads, sewer systems, and other necessary infrastructure. These data can also be used for projecting traffic congestion or crime rates.

Fiscal Impact Analysis

Suppose your town is considering a large residential development project. Since the building of new homes will create increased demands for public services and infrastructure, you would like to prepare the town for such a change. The housing development will create demands for new schools, health care facilities, shopping areas, new roads, sewage systems, energy, pollution control, public transportation, and police and fire protection. At the same time, it will create new employment and generate local taxes, such as those on property and sales. To plan for such extensive changes, you must forecast the impact of this new development. **Fiscal impact analysis** is particularly suited for these kinds of predictions.³

Let us explain this analysis with an example. A developer is planning to build 500 new single-family housing units, of which 40 percent will have two bedrooms, 35 percent will have three bedrooms, and the remaining 25 percent will have four bedrooms. The houses will sell for \$90,000, \$110,000, and \$150,000, respectively. The demographic profile of your town indicates that on average, each two-bedroom house is expected to accommodate 2.354 occupants with 0.148 children. The figures for three- and four-bedroom units are 2.512 occupants with 0.614 children and 2.689 residents with 1.302 children. From this information, you can find the projected demographic profile for the newly developed community (see Table 10.1).

As you can see, using similar logic we can generate a wide range of forecasts, from employment and income generation to the need for additional police, fire, and city employees. The question then becomes, how do you get the necessary numbers

TABLE 10.1 Projected Demographic Profile of New Community

	No. of units	No. of residents	No. of children	Property tax revenue (at 1 percent of sales price)
Two-bedroom units	(500×0.40) 200	(200×2.354) 471	(200×0.148) 30	$(200 \times \$900)$ \$180,000
Three-bedroom units	(500×0.35) 175	(175×2.512) 440	(175×0.614) 107	$(175 \times \$1,100)$ \$192,500
Four-bedroom units	(500×0.25) 125	(125×2.689) 336	(125×1.302) 163	$(125 \times \$1,500)$ \$187,500
Total	500	1,247	300	\$560,000

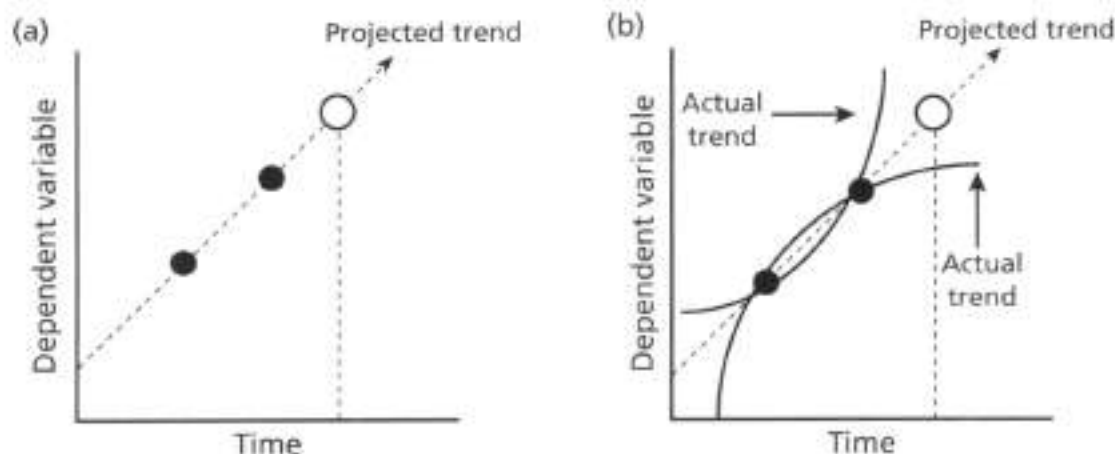
(the “multipliers”) to estimate the future demand, revenues, and costs? There are several methods for getting these numbers:

- The **per capita multiplier method** is the most popular among all others for estimating the basis for future projection because it can be done quickly and relatively inexpensively. This method assumes that the resource utilization pattern of tomorrow will not be any different from today's. For example, to estimate the day care needs of a new community, we can look at the citywide estimate of the average number of children of appropriate age per household. Then, by multiplying that number by the number of new homes, we can calculate the expected number of children in the neighborhood. Since only a fraction of these children will require day care facilities, we can consult other sources for information on the average day care requirements of children from similar social and economic backgrounds. Moreover, using this method, we can estimate such diverse aspects as the number of cars, pollutant emissions, and traffic congestion. If the current average occupant of a three-bedroom house owns 1.98 cars, this pattern will remain unchanged in the next two years, when this hypothetical project is scheduled to be completed.
- The **case study method** involves interviewing municipal department heads, school administrators, and other experts in the field. The premise behind such a method is the assumption that these are the individuals who know the conditions best. Therefore, without the analyst's getting data on per capita resource consumption (as in the previous case), we can get the experts' assessments of the relevant multipliers.
- The **comparable city method** looks at an analogous project and can be used profitably as long as the two projects are similar in nature.

Problems of Single-Factor Analysis

The principal advantage of the methods discussed above is that they are relatively inexpensive and provide quick estimates of the future course of events. As a result, these techniques remain the most commonly used in the area of public sector analysis. We frequently come across estimates derived by these methods in policy debates about crime rates, population growth, and the number of drug addicts. The problem with such estimates is the implicit assumption that whatever happened during the previous period will continue unchanged. However, in life, relationships hardly remain unaltered over time. If we project the rate of AIDS growth in the United States on the basis of the past five years, we can see that there is already some empirical evidence to suggest that the disease's rate of growth may have slowed down, at least among some target groups, as a result of greater public awareness through education and media exposure. Therefore, the reality of the future may not turn out to be as dire as predicted by our method.

Figure 10.2 illustrates the predicaments of a single-factor projection. In Figure 10.2 (a) the two black dots represent two known points in history. On the basis of this meager knowledge, we are predicting the linear trend shown by the arrow. However, an infinite number of nonlinear trends can be drawn through two points. In Figure 10.2 (b), I have drawn two such nonlinear trends. If either of these is the actual trend, a projection based on a straight-line trend will cause serious errors. (For an example of how San Diego County used single-factor projection to estimate the cost of illegal immigration, see “A Case in Point.”)

FIGURE 10.2 Problems of Single-Factor Projection

A CASE IN POINT

Single-Factor Analysis: The Fiscal Impact of Illegal Immigration in San Diego County

Because of the various problems associated with single-factor projections, they are always highly controversial. Yet when no reliable information is available, they can effectively serve as the starting points of policy debate and eventual policy formation. Since these techniques are extremely common in the public sector decision-making process, let me give you an example.⁴

Located on the border with Mexico, San Diego County faces the problem of large numbers of illegal immigrants, mostly from Mexico and Central American countries. Nearly all of these people are extremely poor, so the county must bear a great deal of the costs of health care, education, and law enforcement. In 1992 the state of California commissioned a study to estimate the number of illegal immigrants in San Diego County. However, the researchers faced the problem that nobody had any information on the number of illegal aliens. Therefore, they used some simple but imaginative methods. They estimated the number of illegal aliens by the following procedure.

During the twelve-month period ending September 30, 1991, the Immigration and Naturalization Service apprehended 540,300 undocumented immigrants in San Diego County. It has been estimated by the San Diego Border Patrol that between one in three and one in five undocumented immigrants are actually apprehended. This implies that 1,080,600 to 2,161,200 undocumented immigrants succeeded in entering San Diego County during the year.⁵ The researchers estimated the cost to the state and county governments for the care of the illegal immigrants by calculating their proportion in the total population served by health care services, the education system, social services, and the criminal justice system, and then multiplying that number by the total expenditure. This was claimed to be the total cost to the state and local governments. Following a similar procedure, the researchers estimated the tax revenue contributions by the undocumented alien population of the region. The total net cost to the state and local governments was estimated to be \$145,921,845.

Discussion Points

1. Why is the count of illegal immigrants problematic? How did this study use single-factor analysis to estimate the costs of illegal immigration to San Diego County?
2. Discuss the accuracy of the study's methodology. Can you suggest a better one instead?

JUDGMENTAL METHODS OF PROJECTION

Frequently in life we encounter situations in which forecasting cannot be performed in a structured way, either because of a lack of knowledge of the past or because the causal linkages are too complex to be quantified properly. In such cases forecasts must be based on special insights and intuition. Let us consider an example.

The passage of a piece of legislation through the U.S. political process is a highly complex affair. After it works its way through a maze of committees, it must pass both chambers of Congress. After its passage, if the president signs it, it becomes law. Disagreement between the House of Representatives and the Senate or a presidential veto can easily derail this process. However, it may be necessary to predict the future of a particular piece of legislation so that those with a stake in the matter can be prepared for a change in the course of action. Or consider the case of policymakers in the State Department who are waiting for a certain development to take place in a foreign country and want a reliable forecast of the situation. They must depend on forecasts based on the intuition or subjective judgment of experts. Out of this necessity, a good number of techniques have been developed to deal with these unstructured forecasting needs.

In the previous section, we attempted to estimate the spread of AIDS with the help of only two data points. We also saw the problems posed by such simple techniques. For example, one must consider the effects of increased awareness of the disease and the development of curative and prophylactic drugs. With such a complex issue, it is entirely possible that no single individual possesses all the necessary information to draw a realistic conclusion. We must depend on judgmental methods of projection based on the collective wisdom of experts. In the following sections I discuss a few of these methods.

The Delphi Technique

The name **Delphi technique** was coined after the famous oracle in Apollo's temple in the ancient Greek city of Delphi, where the oracle (in fact, the priests hiding behind it) used to forecast the future of the devotees. The Delphi technique is an important subjective predictive tool that was developed in 1948 by researchers at the RAND Corporation and since the mid-1950s has seen wide use in many countries around the world. Like most of the other techniques in the field of operations research, the Delphi technique owes its origins to attempts to solve problems of military strategy systematically.⁶

The Delphi technique brings a systematic, unbiased reasoning process into subjective group forecasting. We may form a panel of experts and let them sit around a table and come to an agreement about what may take place in the future. However, decisions made by a group may suffer from several sources of bias. If there is a well-known authority in the panel, the lesser members may become intimidated. If the rest of the panel is dispassionate about an issue, they may be swayed by one individual with a strong personality who approaches the issue with a particularly strong opinion. Also, research shows that most people are victims of **groupthink**.⁷ That is, most of us loathe being the odd person out, with a different point of view from the rest of the group. As a result, more cogent points with differing points of view may never be raised in a group discussion.⁸

One way of solving this problem is to ask the experts to forecast independently. In this case, the problem of succumbing to groupthink will be averted, but if the experts have conflicting conclusions, we will have no way of achieving a consensus. Therefore, the Delphi technique was developed to find a happy medium between preserving the individuality of opinion and a synthesis of ideas. It is based on four principles:

1. *Anonymity.* Individual anonymity is achieved through the strict physical separation of panel members. In some cases, even the names of the members should be kept secret from one another.
2. *Iteration.* The judgments of the panel members are summarized and circulated so that group members can modify their original positions. Each round of individual deliberation initiated by information on others' opinions is called an iteration. During the entire process, there may be two or three such iterations.
3. *Distribution of statistical summary.* The individual responses are tabulated, and the measures of their central tendencies and dispersion are provided to the members. To eliminate the extremes, often the median value is presented for the measure of central tendency. For dispersion, a range of measures is usually provided. The members may also receive detailed graphs and charts specifying the shapes of the distribution of responses.
4. *Group consensus.* Finally, on the basis of this process of iteration and feedback, efforts are made to achieve a group consensus on the issue.

The Delphi technique was created primarily for forecasting technical information from a largely homogeneous group of experts. The overall homogeneity in value creates a strong central tendency for the distribution in the forecast values (like a bell-shaped curve). In such a situation, the mean value is a fair representation of the group's judgment on a particular issue. However, when it comes to forecasting a sociological phenomenon, such unimodality of distribution of opinion may not exist. Let us consider, for example, an emotionally charged issue: the future of race relations in the United States. A group of experts assembled to discuss the matter are likely to reflect diversely held value positions and strongly disagree. The group may not be able to form a consensus. In such cases, a slightly different method, called the **policy Delphi technique**, may be more appropriate. This technique starts with the initial assumption that the experts are not homogeneous in their points of view. In fact, the panel members may not even be experts but instead be individuals who represent various interest groups. Therefore, for the policy Delphi technique, the original steps are modified to reflect the changed reality:

1. *Selective anonymity.* It is recognized that there will be subjectivity in arguments on the basis of interest or value positions. Therefore, the participants are frequently kept anonymous only during the initial stage of discussion. After everybody has a chance to state his or her view, the issue may be debated openly in the subsequent iterations.
2. *Informed multiple advocacy.* Unlike the original Delphi technique, this method directs that the panel members be chosen not for their expertise but for their special

interest in or position of advocacy on the matter. The panel considering policy options on how to contain outbursts of racial or ethnic hostility may include conservative advocates of strict law and order, liberal advocates of social reform, and members of opposing ethnic groups.

3. *Multimodal response.* Since opinions are likely to reflect the multimodal distribution of opinion of such a panel, the statistical summary to be provided to the members for the subsequent iterations may not include an attempt to find the central tendency. Instead, the summary may simply provide, as accurately as possible, a picture of the multipolar distribution of opinion.

4. *Structured conflict.* The original Delphi technique depends on the convergence of views, but the policy Delphi technique is built around conflict. In a contentious world, it is often helpful to be able to define opposing points of view clearly. Therefore, policy Delphi does not always aim at resolution, and sometimes it shows a final unbridgeable gap between parties.

Both Delphi techniques seem deceptively simple. For their successful use, you must follow the same path of structured reasoning as we discussed in chapter 5. The steps are outlined here.

Define the problem. The success of this process depends on the clear definition of an issue. For example, an agency faces a probable cut in funding. Before it assembles a panel, it must decide the perimeter of the issue: should it attempt to forecast the amount of money available for the next fiscal year, or should it tackle the question of specific cuts corresponding to certain levels of funding? An ill-defined issue can easily cause confusion and cost the organization a great deal in wasted effort, time, money, and morale. The proper definition of the issue is even more critical for a policy Delphi analysis, because social issues are likely to be far more complex than a technical problem facing an organization.

Choose the right panel. Choosing the right panel is equally critical for the success of the Delphi technique. Hard thinking must precede the selection of the panel members. Often the individuals designing a Delphi analysis may not have adequate knowledge of the important persons relevant to the issue. William Dunn suggests a practical solution to this problem.⁹ Frequently, the planners are at least able to name the most influential figure in the debate. They may ask this individual to identify the person with whom he or she agrees most closely and another with whom he or she disagrees most vehemently. By asking these individuals the same question, the planners are well on the way to selecting an entire panel that shows a full range of opinion.

Develop the first-round questionnaire. The success of Delphi depends on the types of questions that are put before the panel, and you, the analyst, must decide what the questions will be in the first and subsequent iterations. Although there are no hard-and-fast rules about developing a questionnaire, you must develop it for the first round with an eye to the next. Suppose the purpose of the exercise is to obtain a forecast, say, of the number of AIDS victims in the next five years. Some very

structured questions regarding the future spread of the disease can start off the discussion. Or the panel can discuss trends in people's attitudes and sexual practices, changing social mores, and the attitude of the administration toward a frank discussion of unsafe sexual and intravenous drug use practices among the target groups and the distribution of prophylactic devices. In this case, the first-round questionnaire can be relatively unstructured and contain a number of open-ended questions. If the questions are not open ended, the answers should be quantified according to some scale.

Analyze first-round results. The results of the first round of questionnaires should be analyzed to determine the position of each panel member. These results should be tabulated, and for each question, the measures of central tendency and dispersion should be calculated. The panel members should have these results available for subsequent rounds. For example, if the question was, "How much do you expect teenage sexual practices to change in view of an increase in awareness campaigns?" and the answers were rated on a five-point scale (5 being significantly changed, 1 being no change at all), the panel members should be given at least the mean, median, standard deviation, and range of distribution of the answers. If you have graphics capability, it may not be a bad idea to show panel members visually the distribution of their answers. This may be a particularly good idea if the members are not expert statisticians.

Develop questionnaires for the subsequent rounds. Comparison of the group results with the individual responses paves the way for further discussion in the Delphi process. If the answers indicate a significant responsiveness of teenage sexual behavior to a concerted ad campaign, a more detailed discussion of this topic can help develop forecasts (and future policies). Also, note that although panel members may not have stated their basic assumptions in the first round, they are allowed to do so in successive rounds. Policy Delphi usually involves three to five rounds, so members have ample opportunity to evaluate one another's arguments in greater detail. These rounds of discussion may cause the members to modify their positions.

Arrange the panel discussion. At the end of the process, a group meeting can allow the panel members to see if a consensus finally emerges through an open, face-to-face discussion. These group discussions can be particularly fruitful, since by now, each member of the panel is thoroughly conversant with the positions, arguments, hypotheses, and logic of others on the panel. Therefore, the group discussion can often take place in an atmosphere of mutual understanding, if not agreement.

Prepare the final report. The last step of a Delphi process is for the analyst to prepare the final report. It should describe the entire process. If an overall agreement appears, the report should mention it, but you should be careful not to ignore minority or extreme positions, if any. If there is no consensus, you must take care to document the diverse points of view and the extent of divergence of opinion. (For a report on how California State University, Long Beach, used the Delphi process to revamp its public affairs program, see "A Case in Point.")



A CASE IN POINT

The Use of the Delphi Technique in Devising a Public Policy Curriculum

I have mentioned the use of the Delphi technique as a forecasting tool. You can extend this and all other techniques of numerical analysis in solving problems far beyond their principal use. For example, the members of the public administration department at California State University, Long Beach, wanted to revise its public policy curriculum, which would fundamentally reorient the direction of the department's program.¹⁰ However, public policy may be taught from many different angles, requiring a wide variety of analytical skills. Since the faculty members could not come to an agreement, they decided to use the Delphi technique to devise the new curriculum. For that, the faculty took the following steps:

1. *Identification of the problem.* The faculty first wanted to know "what practicing public managers, academics, students, and alumni thought MPA [master of public administration] graduates should know and be able to do, and second, to rate or rank the relative importance of those knowledge and skill components." They also wanted to identify "which components should be taught in required courses and which in elective."¹¹
2. *Identification of the monitor/team.* Three faculty members with the required skill and knowledge of the Delphi technique were selected to make up the research team.
3. *Identification of the sample.* The team identified the stakeholders (the students, alumni, faculty members, and public managers), from whom they chose more than 100 participants for the first round. The participants were required to have access to e-mail. This requirement was not seen as a bias in the study, since nearly 85 percent of the students worked for public agencies and had e-mail addresses. High percentages of alumni, faculty members, and public managers also had e-mail access.
4. *Round 1.* The participants received two open-ended questions: What should MPA graduates know, and what skills should they have?
5. *Round 2.* The research team tabulated the responses under the headings "knowledge" and "skills" and sent them back to the participants with the request to add, delete, or otherwise modify the lists.
6. *Round 3.* The team compiled complete lists under "knowledge" and "skills" and sent them back to the participants, asking them to award weights from 0 to 10, with 10 being the most important to the MPA students.
7. *Round 4.* The team members tabulated the responses for mean and standard deviation and ranked them according to their relative importance. Finally, they sent the results back to the participants and asked them to assign "core" or "elective" to each of the categories of "knowledge" and "skills."

8. *Sharing of information.* The faculty shared the gathered information and from it devised seven required and five elective courses.

Here is the list of the ten most important topics in the "knowledge" and "skills" categories, ranked from most to least important:

Knowledge	Standard deviation	Mean
1. Economics, politics and markets, microeconomic theory	9.7	0.55
2. Administrative theory, theory of the state	9.5	0.70
3. Research methods	9.5	0.70
4. Policy analysis	9.0	1.40
5. Strategic and comprehensive planning	9.0	1.40
6. Statistics	9.0	1.40
7. Financial management, politics of financial management	9.0	1.40
8. Program evaluation	9.0	1.40
9. Budgeting	8.5	2.10
10. Intergovernmental relations	8.0	1.00

Skills	Standard deviation	Mean
1. Writing skills	9.4	0.55
2. Analytical skills and techniques	9.0	1.00
3. Policy analysis skills	8.6	1.34
4. Leadership skills	8.4	1.14
5. Oral presentation skills	8.4	3.50
6. Finance and budgeting skills	8.2	1.48
7. Political awareness and interaction	8.0	1.58
8. Interpersonal management/behavioral skills	8.0	1.87
9. Cultural competency/diversity management	7.8	1.30
10. Listening skills	7.8	2.95

Discussion Points

1. Discuss the relative importance of the top ten "knowledge" and "skills" requirements, particularly in light of their mean and standard deviation scores.
2. Can you think of other areas in which the Delphi technique can be profitably used?

The Feasibility Assessment Technique

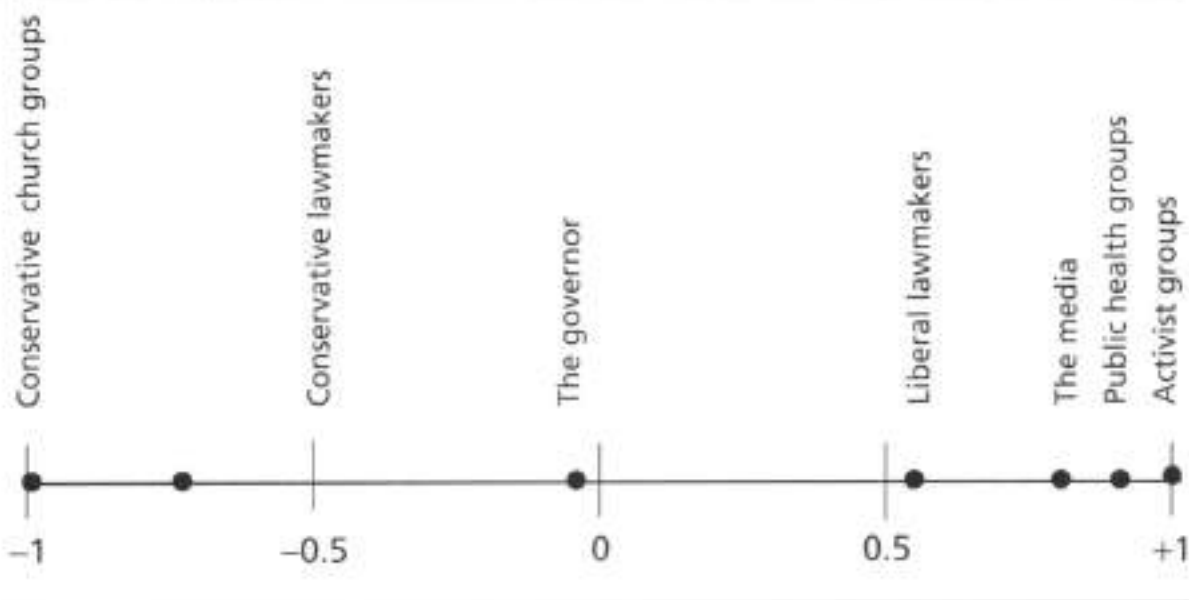
The **feasibility assessment technique** (FAT) is a commonly used judgmental method of projection. This method is particularly useful in forecasting the outcome of a contentious issue, fought by a number of interested parties. Therefore, FAT has found wide application in the forecasting of political, economic, military, and institutional outcomes of a conflict. It is a versatile technique and can be used at any phase of the policy-making process. It can be used for predicting which issues will come to the forefront, what the legislative outcomes will be, or how a policy will be implemented.

For example, suppose we are trying to forecast the outcome of a bill before a state legislature to fund the distribution of AIDS information to high school students. To forecast the outcome of this debate, we need to make a list of the interested parties who play an active role in determining the outcome of the bill. We can make a list of the major players as follows: the governor, the liberal lawmakers, the media, the conservative lawmakers, the public health groups, the conservative Christian church groups, and the gay activist groups. Let us assume that the governor is a moderate conservative who shows mild opposition to the bill. The liberal lawmakers strongly support the measure, and the conservative lawmakers are solidly opposed to it. The public health groups support the issue. Finally, the church groups are vehemently opposed to the public funding of explicit sexual education, and the gay activists are equally visceral in their support for the bill. In FAT terminology, the relative position of the "player" groups (those who exert some power and influence over the policy outcome) is called the "issue position."

An analyst (or a group of analysts) assesses the relative issue position by estimating the probability of support by each player group. This probability measure is assigned to each group, and it varies between +1 and -1. If a group is certain to support the issue, its issue position will have a value of 1. If it is certain to oppose it, its issue position value is assessed as -1. If the group is likely to be indifferent, its value will be 0. It is safe to assume that the gay activist groups should be assigned a value of +1, and the conservative church groups should be given a value of -1. Let us assume that the support of the public health groups is 0.9, the media 0.8, liberal lawmakers 0.6, the governor -0.1, and the conservative lawmakers -0.7. The relative positions are shown in Table 10.2. Note that we are assigning values on the basis of our subjective assessments. In a class exercise, you can simply make your estimations. However, a great deal of serious thinking has gone into "mental modeling," by which an expert in the field can arrive at a subjective

TABLE 10.2 Relative Issue Positions

Group	Relative position
Gay activist groups	+1.0
Public health groups	+0.9
Media	+0.8
Liberal lawmakers	+0.6
Governor	-0.1
Conservative lawmakers	-0.7
Conservative church groups	-1.0

FIGURE 10.3**Relative Issue Positions on Public Funding of AIDS Information**

assessment. Analysts must conduct these subjective forecasting techniques with the utmost care because of the influence of a great number of possible biases. I have mentioned some of these biases at the end of this chapter.¹²

Looking at Figure 10.3, you might think that because more groups with passion appear to the left of the spectrum (the right-hand side of the diagram), the issue is certain to be decided in favor of funding. Reality, however, is more complicated than that. To begin with, not every player has equal influence over the outcome of the debate. For the next step, we would want to measure the players' available resources on the issue. The resources within the disposal of each group would include prestige, legitimacy, money, time, administrative capabilities, and communication capabilities. The available resources are measured within the range of 1 (having a great deal of resources to bear) and 0 (having no resources whatsoever). We may hypothetically assign values to our lineup as shown in Table 10.3.

TABLE 10.3**Availability of Resources**

Group	Resource availability
Governor	0.90
Liberal lawmakers	0.40
Conservative lawmakers	0.50
Media	0.80
Conservative church groups	0.30
Gay activist groups	0.10
Public health groups	0.05

TABLE 10.4 Potential for Policy Influence

Group	Issue position (a)	Available resources (b)	Potential for policy influence (a × b)
Governor	-0.10	0.90	-0.09
Liberal lawmakers	0.60	0.40	0.24
Conservative lawmakers	-0.70	0.50	-0.35
Media	0.80	0.80	0.64
Conservative church groups	-1.00	0.30	-0.30
Gay activist groups	1.00	0.10	0.10
Public health groups	0.90	0.05	0.045
Total			+0.285

The potential influence of each player can now be calculated by multiplying the issue position of each group by its total available resources. Thus, we can derive Table 10.4. The total potential of policy influences predicts a positive outcome for the measure. However, the potential does not often foretell the actual outcome. Every player faces a slate of issues it considers to be vital to its mission. The governor has many agendas, of which fighting the funding of the AIDS information project is one. Facing an issue, the players must decide on the relative importance of the issue in relation to their other obligations. In other words, this particular issue, like all others facing each group, must have the group's commitment to invest a percentage of its resources. This commitment is called the "ranking of resources." Let us assume that the players have decided to allocate their total available resources as shown in Table 10.5.

In other words, the liberal lawmakers are willing to commit 15 percent of their resources, and the gay activist groups are estimated to spend 80 percent of their resources, whereas the largely disinterested media are expected to commit no more than 2 percent in promoting this measure. The governor has decided to spend no more than 10 percent of his resources in fighting this measure, but the conservative lawmakers have made this issue a hallmark of their conservative agenda and are expected to spend 30 percent of their available resources. Similarly, the conservative church groups are expected to commit 90 percent of their resources to fight public funding of what they

TABLE 10.5 Ranking of Resources

Group	Resource ranking
Governor	0.10
Liberal lawmakers	0.15
Conservative lawmakers	0.30
Media	0.02
Conservative church groups	0.90
Gay activist groups	0.80
Public health groups	0.60

TABLE 10.6 Calculated Policy Influence

Group	Potential for policy influence (a)	Ranking of resources (b)	Feasibility score (a × b)
Opponents			
Governor	-0.09	0.10	-0.0090
Conservative lawmakers	-0.35	0.30	-0.1050
Conservative church groups	-0.30	0.90	-0.2700
Subtotal			-0.3840
Proponents			
Liberal lawmakers	0.240	0.15	0.0360
Media	0.640	0.02	0.0128
Gay activist groups	1.000	0.80	0.0800
Public health groups	0.045	0.60	0.0270
Subtotal			0.1558
Total			-0.2282

consider to be offensive. This commitment factor multiplied by the total potential determines the outcome. Thus, we may construct Table 10.6.

Table 10.6 gives us the total feasibility score: the opponents of the measure will have more support than will the proponents (-0.3840 as opposed to 0.1558). Therefore, despite popular support (more groups in favor), the measure will be defeated.

You may find the conclusion of this hypothetical study somewhat surprising, especially in view of the initial assessment based on Figure 10.3 and Table 10.4. In fact, the ultimate outcome of any public policy depends on the relative issue position and the fraction of the total resources that the player groups are willing to invest in achieving a favorable outcome. You may notice that although the governor is the most resourceful person in this debate, his reluctance to invest a great deal of resources reduces him to the position of a minor player. In contrast, by combining their total resources and a stronger determination, the coalition between the conservative lawmakers and the church groups becomes a formidable force in stopping public funding. This technique explains very well many different social and political events. Consider, for example, the impact of the so-called Moral Majority, a coalition of right-wing religious groups and conservative politicians put together by the Reverend Jerry Falwell in the late 1970s and early 1980s. Although not supported by the majority of the American public, this group had a profound impact on the course of American politics that went far beyond its numerical strength.

The steps of FAT can be summarized as follows:

1. Identify the issue.
2. Identify the player groups.
3. Estimate the issue positions of the groups.
4. Estimate the available resources for each group.
5. Estimate the resource rank within each group.
6. Calculate the feasibility assessment index.

The Expected Utility Model

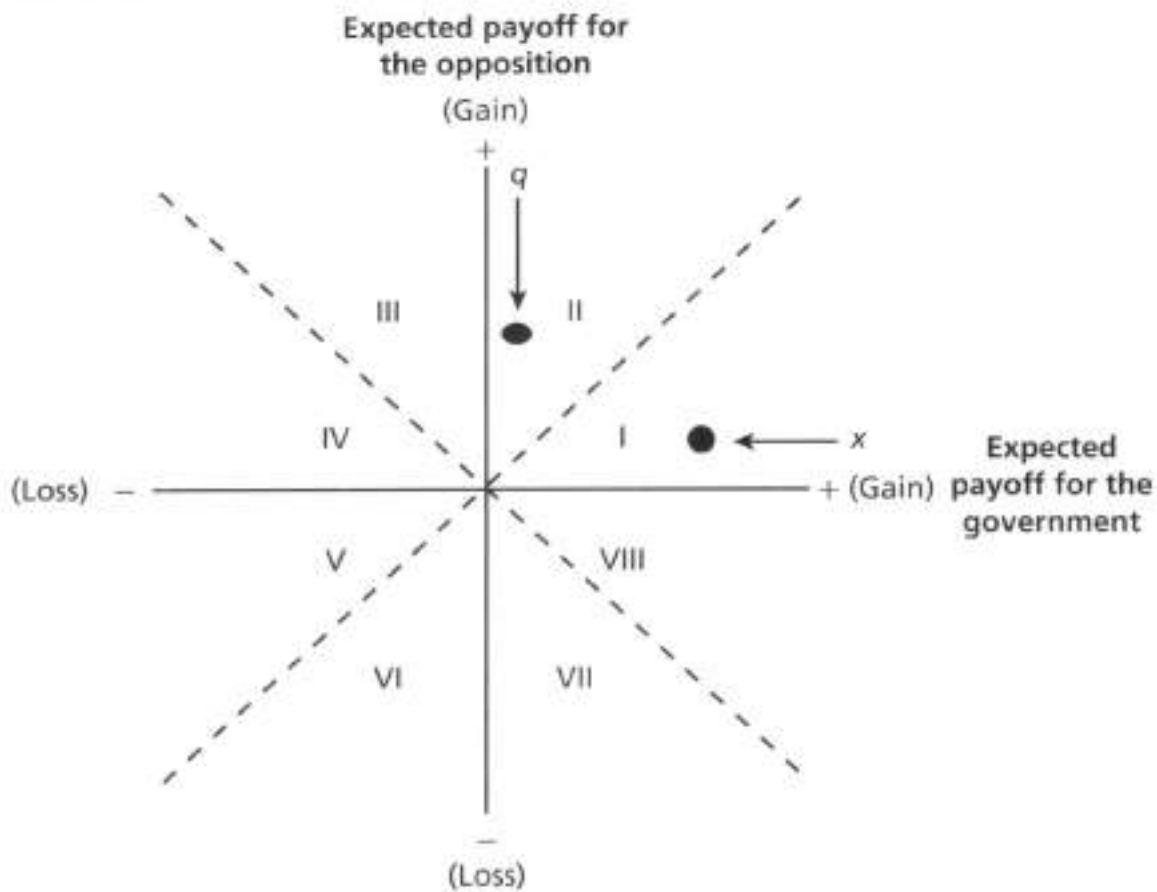
Some analysts have found a variation of the feasibility assessment technique to be a useful forecasting tool for predicting the outcomes of an incredible variety of social phenomena, from international relations to banking regulations. Bruce Bueno de Mesquita of the Hoover Institution at Stanford University and his associates are at the forefront of such predictive efforts, which they call the **expected utility model**.¹³ Their methodology for forecasting is far too complex to be discussed in this book, but I can give you the basic idea behind these prediction methodologies.

Suppose we are forecasting the probability of a change in a certain government policy. By scanning the political landscape, we can pick out the major players in the game. They may then be classified as proponents of a change or opponents of a change. Let us assume that the government and its allies do not want any change and the opposition groups do. Those groups that are proposing a change are inviting a confrontation with those that prefer the status quo. Each group recognizes that just like investing in a risky project, the investment of resources to fight a rival has its own risks. A loss might cause an embarrassment and expose the group's vulnerability to its foes. In contrast, a win will bring highly desired spoils. Therefore, a group's strategic move in determining whether to confront its opponent and how much to invest in the process will reflect its expectations about the future. The following example can help show the relative positions of the proponents and the opponents of a policy change.

Suppose opposition groups propose a change in government policy (for example, in handgun control, or increased funding for urban renewal, or conservation of open space), whereas the government and its supporters oppose such a change. The expected payoffs of the two groups can be shown with the help of a diagram. In Figure 10.4 the quadrants created by the intersection of the two straight lines have been subdivided into octants. These octants are marked with Roman numerals. The expected payoffs of the two groups are plotted on a Cartesian plane. In the northeastern quadrant, comprising octants I and II, each contestant expects to gain by confronting the other. If both contenders feel that they can win in a confrontation, the chances of open confrontation are extremely high. However, if the government's expectations of the outcome of a confrontation fall in octant I, at point *x*, then the government expects to gain more than what the opposition parties would gain. Therefore, we can expect the government to take an aggressive posture and start a confrontation. In an authoritarian regime, this confrontation may take the form of cracking down on the opposition. In a democratic system, it may be a presidential veto.

In contrast, if the government's expectations of the outcome of a confrontation fall on a point such as *q* on octant II, then it would expect to win, but its win would be less than that of its rival. As a result, if mutual expectation patterns fall in this region, we can expect a role reversal, wherein the opposition parties more aggressively seek confrontation with the government. However, since both parties expect to win, the chances of a confrontation are extremely high. In an international context, disputes falling within these two octants have escalated into war about 90 percent of the time.¹⁴

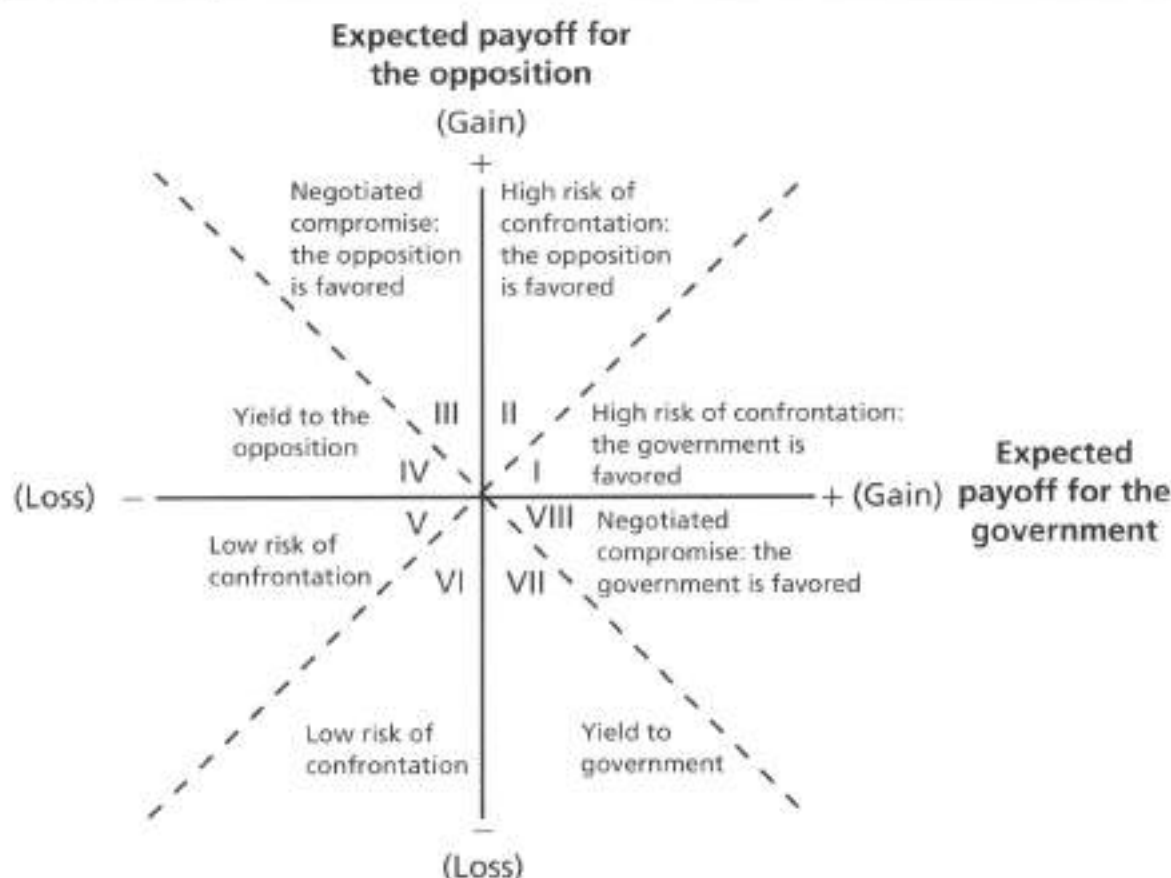
In the northwestern quadrant (octants III and IV), the opposition expects to win and the government expects to lose. In octant III, although the challenger expects to win, its gains will be smaller than the other's loss. Therefore, seeing the prospect of a relatively small loss, the government will sit down with the opposition and negotiate a compromise that gives the opposition an edge.

FIGURE 10.4 Expected Payoff of the Government and Its Opposition

In octant IV, in contrast, the government is expected to suffer a heavier loss than what the opposition would gain. If the government's expectations match this kind of a pattern, we can expect it to accede to the demands of the opposition. In 1990, as part of a wide-ranging budget compromise, President George H.W. Bush decided to go along with the Democrats in increasing taxes. At that time the president decided he had more to lose by opposing a compromise. In view of his earlier unequivocal campaign pledge of not increasing taxes ("Read my lips, no new taxes"), however, his action became a liability and came back to haunt him later.

In the southwestern quadrant, both parties expect to lose. You may note that in octant V, the government's losses are expected to be greater than the losses of the opposition. The situation is reversed in octant VI. Facing the prospect of a lose-lose situation, neither party shows much enthusiasm for a head-on confrontation.

Finally, the southeastern quadrant is the mirror image of the northwestern quadrant. That is, in this case, the government is expected to win in confrontation, and the opposition expects to lose. When the condition of octant VII prevails, the government perceives that it is making a legitimate demand on the opposition, and it has a good deal to gain from its position; but more important, the opposition has a good deal more to lose. Situations such as these prove to turn out peacefully because the opposition is effectively shut off with the prospect of a heavy loss. In the spring of 1984, President Ferdinand Marcos of the Philippines called for elections. President Marcos often used

FIGURE 10.5 Dynamics of the Policy Outcome

the results of highly corrupt elections to gain political legitimacy for his regime. The widespread disenchantment with his rule prompted many observers to predict organized opposition to Marcos or a call for the boycott of the elections and even violence. However, the leading opposition groups did not expect to defeat Marcos and expected to lose far more than the loss they could inflict on his regime by opposing him in the election. On the basis of their perception, Bueno de Mesquita correctly predicted no serious challenge to Marcos's presidency and a relatively uneventful election. I show the expected outcomes of confrontation in Figure 10.5.

Origin of the Numbers

The numbers in a FAT model or the estimation of the expectations of the two rivals in an expected utility model come from the judgment of those who are experts in the area. The expected utility model requires the identification of the major players and the estimation of their expectations. For these estimations, similar to those in the FAT model, you can use the expertise of those who have direct knowledge of these groups. Having ascertained the groups' positions, you may first plot them in their respective octants to see if a pattern emerges. These results will tell you whether you can expect a confrontation.

Although I have provided examples from international relations, a versatile technique such as the expected utility model can be used to forecast many different outcomes in contentious situations facing all levels of government. Let us suppose that during a police chase in your town, a police cruiser crashed into the car of an innocent citizen, injuring her. The risk management division of your local government is interested in settling the case. However, before offering a settlement amount, the risk manager wants to know if the injured party would want a lengthy and extremely expensive court case. The expected utility model can help you develop a scenario that will justly compensate the victim for her suffering and yet save a good deal of the city's tax dollars.

Shortcomings of the Judgmental Methods

Often in life, issues are so complex that it is extremely difficult to capture all their dimensions within the confines of structured techniques of analysis and forecasting. In addition, in real life you will encounter situations in which there is no history to depend on for predicting the future. In such cases, judgmental methods give us extremely useful alternatives. Time and again, the various subjective methods of prediction have proved their usefulness. However, the very sources of their strength in one situation become their liability in another.

Although the techniques of judgmental methods offer more flexibility than do the other, more rigorous statistical methods (discussed in the following chapters), they still must operate within some structure. In analyses of group interaction (say, between the government and the opposition), we assume that all groups are making their decisions at the same time yet independently. It stands to reason, however, that they are interdependent. Also, group interactions are part of a dynamic process in which groups communicate with one another, compromise, and form or break coalitions. Even the most flexible open-ended forecasting method cannot accommodate many of these complex behaviors.

However, the more serious problem with the judgmental methods is that because there is no systematic way of discussing the assumptions and arguments of the "experts," preconceived ideas, prejudices, unarticulated agendas, self-interest, or other psychological pitfalls and cognitive limitations frequently contaminate their judgments. Few experts (in and out of the government) predicted the first war between the United States and Iraq in 1991. Even at the last minute, most were willing to dismiss the threats and counterthreats (for example, Saddam Hussein's apocalyptic promise that it would be the "mother of all wars") as simple bravado, posturing, or face-saving gestures. When President George H. W. Bush ordered Iraq to get out of Kuwait or face war, he was not bluffing. And when Hussein, facing incredible odds, said that his army would never leave Kuwait of its own accord, he meant it. Yet most experts saw the prospect of this strange war as so far-fetched that they refused to believe the resolve of the two contenders almost until the first bullet was fired. Similarly, because of these psychological impediments, the Central Intelligence Agency was caught off guard by the sudden collapse of the Soviet Union. Having poked and prodded, studied and analyzed the Soviet system for decades, agency experts completely failed to understand the fundamental fragility of the system.

KEY WORDS

Assumption of continuity (p. 225)
Case study method (p. 230)
Causal connections (p. 225)
Comparable city method (p. 230)
Delphi technique (p. 232)
Dependent variables (p. 225)
Expected utility model (p. 242)
Feasibility assessment technique (p. 238)
Fiscal impact analysis (p. 229)

Groupthink (p. 232)
Independent variables (p. 225)
Judgmental methods of projection (p. 227)
Per capita multiplier method (p. 230)
Policy Delphi technique (p. 233)
Prediction (p. 225)
Projection (p. 225)
Single-factor projection (p. 227)
Trends (p. 225)

EXERCISES

1. Consider the following facts in your community and explain how you would go about forecasting them for the next five years:
 - a. The number of homeless people
 - b. The number of child abuse cases
 - c. The number of school-age children
 - d. The number of violent crimesExplain your data needs, point out the possible sources, and choose a projection model.
2. Suppose the indigent elderly population that depends on public assistance is growing at a rate of 4 percent per year in your county. At present there are 3,500 individuals in this category. How many such people would you expect to see three years from now? Assuming that it costs the county \$2,300 per person for health care, estimate the total cost of indigent health care for elderly people in your county.
3. Explain the problems of single-factor projections. Despite these problems, single-factor projection remains one of the most commonly used techniques of projection. Taking a real-life example, account for the widespread use of single-factor projection in public policy analysis.
4. Explain the process of the policy Delphi technique. What are its strengths and shortcomings?
5. The city of Masters, Pennsylvania, faces a controversial issue. The marshland adjacent to a prosperous neighborhood is not being used. A developer has submitted a proposal to make it into a golf course. However, the marshland is the habitat for migratory birds. Hence, the project is opposed by powerful environmental groups. The conservative, probusiness council members support the project. This is a divisive issue, and the mayor has expressed her mild opposition to the project, whereas the liberal members of the city council are in vehement opposition. The table shows the relative positions of the various parties, their available resources, and the ranking of resources. Predict the outcome of the debate.

Group	Issue position	Ranking of resources	Available resources
Mayor	-0.20	0.30	0.80
Developer	11.00	0.90	0.30
Conservative city council members	0.85	0.60	0.40
Liberal city council members	-0.65	0.80	0.50
Environmental groups	-0.90	0.90	0.90
Chamber of commerce	10.50	0.40	0.10

6. Take any current controversial issue facing your community, state, or the nation, such as gun control, abortion, or the amount of national defense funding. Then, working with others in a group project, predict the policy outcome using the feasibility assessment technique. How accurate do you think your predictions are? What are the major weaknesses of your predictions? Would you feel comfortable using these techniques in a real-life situation? Explain. (Suggestion: The size of the group is often crucial to the success of such a project. A group size of about five is best. A larger group can be broken into several groups, each one attempting to forecast the same issue or different ones.)
7. Suppose you are going to forecast the outcome of an upcoming sporting event. As a group, discuss and write down the factors (quality of the players, past performances of the two teams, and so forth) that may help you in your prediction. Then, using the judgmental method, develop a forecast. Compare your forecast with the actual outcome of the game.

Notes

1. Those of you who have hand calculators with a button that says [y^x] can be spared the task of having to multiply 1.09 ten times. The [y^x] key raises a number to the desired exponent. To calculate this number, you need to (1) enter 1.09, (2) hit the [y^x] button followed by (3) the number 10, (4) press the [=] button to obtain the value for the expression $(1 \times 0.09)^{10}$. By multiplying this number by the initial number of cases (P_0), you can estimate the value of P_{10} .
2. To arrive at this number, we write equation 10.5 as follows:

$$\sum_1^{10} P_n = \frac{(1 + 0.09)^{10+1} - 1}{0.09} \times 150$$

3. For a detailed discussion of fiscal impact analysis, see Robert W. Burchell and David Listokin, *The Fiscal Impact Handbook: Estimating Local Costs and Revenues of Land Development* (New Brunswick, N.J.: Rutgers University Press, 1978).
4. This study generated a lively controversy on the cost of illegal immigration in the region. Without taking any side in the controversy, I am simply reporting the results as an example of single-factor projection when no other data are available.

5. Auditor General of California, "A Fiscal Impact Analysis of the Undocumented Immigrants Residing in San Diego County," Report C-126, August 1992, 11.
6. For a historical account of the development of the Delphi technique, see Harold Sackman, *Delphi Critique: Expert Opinion, Forecasting, and Group Process* (Lexington, Mass.: Lexington Books, 1975). See also Juri Pill, "The Delphi Method: Substance, Contexts, a Critique, and an Annotated Bibliography," *Socio-Economic Planning Science* 5 (1971): 57-71. However, for an excellent discussion of Delphi and other techniques of subjective decision making, see William N. Dunn, *Public Policy Analysis: An Introduction* (Englewood Cliffs, N.J.: Prentice Hall, 1981).
7. The term *groupthink* was made popular by the psychologist Irving Janis. For a more detailed explanation and many more examples, see Irving Janis, *Groupthink: Psychological Studies of Policy Decisions and Fiascoes* (Boston: Houghton Mifflin, 1982), a revised and enlarged edition of Irving Janis, *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes* (Boston: Houghton Mifflin, 1972).
8. People often choose not to express their "true preferences" in public because of peer pressure or the fear of social ostracism. If the need to be "politically correct" overwhelms the need to be truthful, your publicly held views will deviate from what you prefer or believe to be the truth. The price of this distortion of a citizen's preferences may be considerable, as public policies based on misconstrued views can lead to inefficiency or social conflict. For an interesting discussion of the need for anonymous discourse for the articulation of genuine preferences, see Timur Kuran, "Mitigating the Tyranny of Public Opinion: Anonymous Discourse and the Ethic of Sincerity," *Constitutional Political Economy* 4 (1993): 41-78.
9. Dunn, *Public Policy Analysis*, 198.
10. Michelle A. Saint-Germain, John W. Ostrowski, and Martha J. Dede, "Oracle in the Ether: Using an E-mail Delphi to Revise an MPA Curriculum," *Journal of Public Affairs Education* 3 (2000): 161-172.
11. *Ibid.*, 163.
12. See, for example, Herbert Simon and Dorothea Simon, "Individual Differences in Solving Physics Problems," in *Children's Thinking*, ed. Robert S. Siegler (Englewood Cliffs, N.J.: Prentice Hall, 1986). See also David Klahr and Kenneth Kotovsky, eds., *Complex Information Processing: The Impact of Herbert A. Simon* (Hillsdale, N.J.: Lawrence Erlbaum, 1989); Vadim D. Glezer, *Vision and Mind: Modeling Mental Functions* (Hillsdale, N.J.: Lawrence Erlbaum, 1995).
13. See, for example, Bruce Bueno de Mesquita, David Newman, and Alvin Pabushka, *Forecasting Political Events: The Future of Hong Kong* (New Haven, Conn.: Yale University Press, 1985).
14. Douglas Beck and Bruce Bueno de Mesquita, "Forecasting Policy Decisions: An Expected Utility Approach," in *Corporate Crisis Management*, ed. Steven J. Andriole (Princeton, N.J.: Petrocelli, 1984).

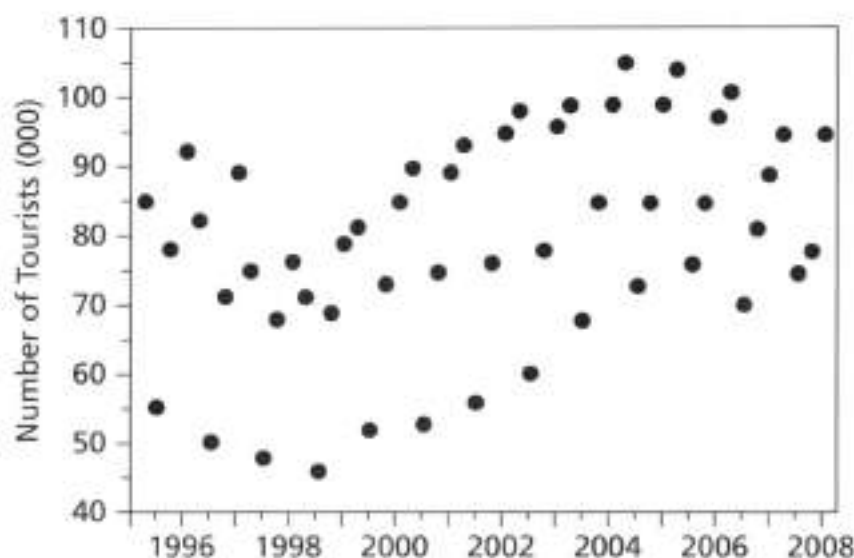
Projection Techniques: Analysis of Historical Data

Forecasting on the basis of projection techniques requires an understanding of the basic pattern of behavior. Without a pattern, variations are totally erratic or random, in which case future behavior cannot be predicted. In the previous chapter we discussed using forecasting methods when faced with a paucity of past information. In such instances, when making projections, we must rely on one or two data points over time or on expert judgment. To make the most of what little historical data we have, we use projection techniques specially suited to situations of data scarcity, such as single-factor projection and judgmental methods of forecasting. However, we select different techniques when we have a long history of behavior to study. In this chapter we examine those techniques that lend themselves to discerning trends in series of data.

To understand the trend of a series (or the direction in which it is heading), we need to plot the data first. Sometimes, though, a scatter plot does not reveal a strong trend. Consider, for example, a hypothetical series of tourists visiting the Pennsylvania city of Masters, which has a strong tourist industry that attracts vacationers primarily during the skiing and summer seasons. The city's chamber of commerce has collected data on the number of tourists quarterly (in three-month periods) for thirteen years, from 1996 to 2008. These data are plotted in Figure 11.1.

THE COMPONENTS OF A DATA SERIES

You may conclude that the scatter plot of the quarterly tourist population looks quite confusing and that you can determine little from it. Often a series contains elements that obscure our vision. In general, a data series tends to head in a general direction, or contains an element of a trend. A series that grows over time is called a **positive trend**, whereas a series that declines is said to have a **negative trend**. Besides

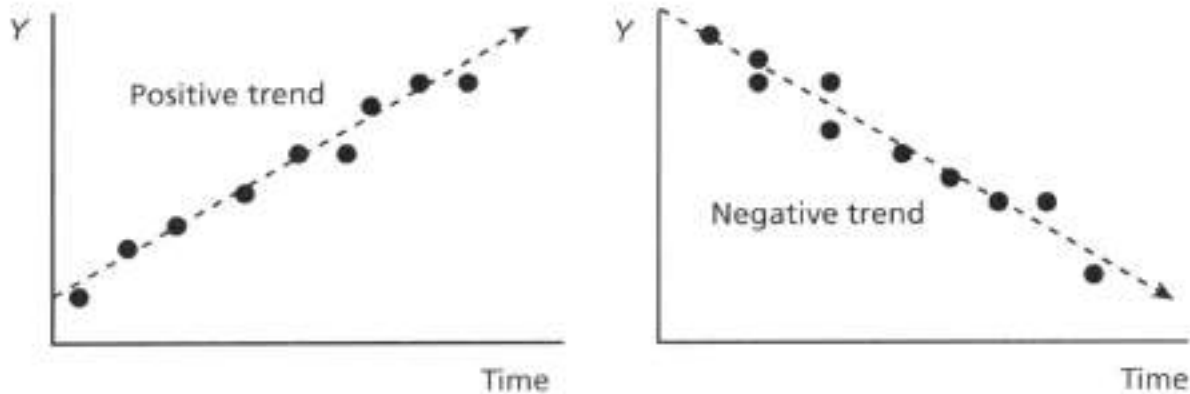
FIGURE 11.1**Number of Tourists in Masters, Pennsylvania (in thousands)**

the trend, a data series may also reflect fluctuations resulting from the effects of seasonality. Furthermore, a data series may exhibit the effects of long-term economic cycles. As the economy expands and contracts during the course of a business cycle, it affects many kinds of economic and social activities. Finally, a series contains the effects of purely random fluctuations, which are caused by factors outside our consideration. Natural calamities (floods, earthquakes, devastating tornadoes, and the like), political events (riots, assassinations of important political persons, the election of a new chief executive or political party, and so forth) or institutional factors (such as changes in government regulations) can have an unpredictable effect on a series. The collapse of the Soviet Union had an extraordinarily important but unforeseen effect on the U.S. defense industry. Since these external factors cannot be factored into a model, their effects are called **random errors**. Therefore, time-series data (data recorded over time) may contain the following:¹

$$\text{Data} = \text{Trend} + \text{Seasonality} + \text{Cyclical Effect} + \text{Random Error}$$

Let us examine these components in greater detail. Figure 11.2 depicts two **linear trend** patterns. Many aggregate social, economic, and demographic data (such as the per capita gross domestic product [GDP] of a country and population growth in the short term) show positive linear growth when plotted over time. However, as shown in Figure 11.2, not all points of observation fall on the trend line. Since these deviations cannot be accounted for within the model, we assume that aberrations reflect the effects of the random error component.

In the case presented in Figure 11.1, the rate of growth does not seem to change over the delineated time period. This trend can be viewed as upward or downward. Despite little peaks and valleys for economic booms and recessions, the GDP per capita, measured in constant dollars, has shown an upward trend over the years in the United States.

FIGURE 11.2 Linear Trend Patterns

But during the 1980s, federal government assistance to state and local governments demonstrated a steady downward trend.

When data values fluctuate around a constant mean, a **horizontal trend** pattern develops. This is characteristic of a series in which there is no trend and the data seem to fluctuate in a random fashion. In technical terms, such a series is called a **stationary series**. In a stationary series, a trend line will go through the mean ($\bar{Y}$). Since there is no seasonality in the human reproductive process, for example, the daily number of births in New York City hospitals within a year will depict such a series. Figure 11.3 is an illustration of a horizontal pattern of fluctuation.

A **seasonal trend** is shown in Figure 11.4. The quarterly data plotted in the figure depict a trend of regular fluctuations of peaks and valleys during the course of a year. Home construction, unemployment, crime, and highway accidents often tend to show this kind of seasonal variability.

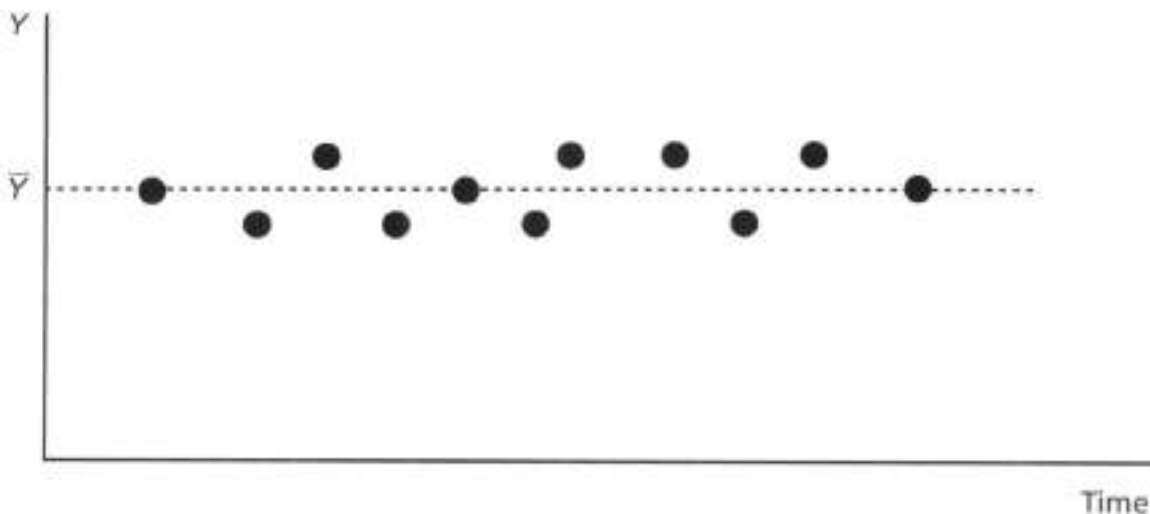
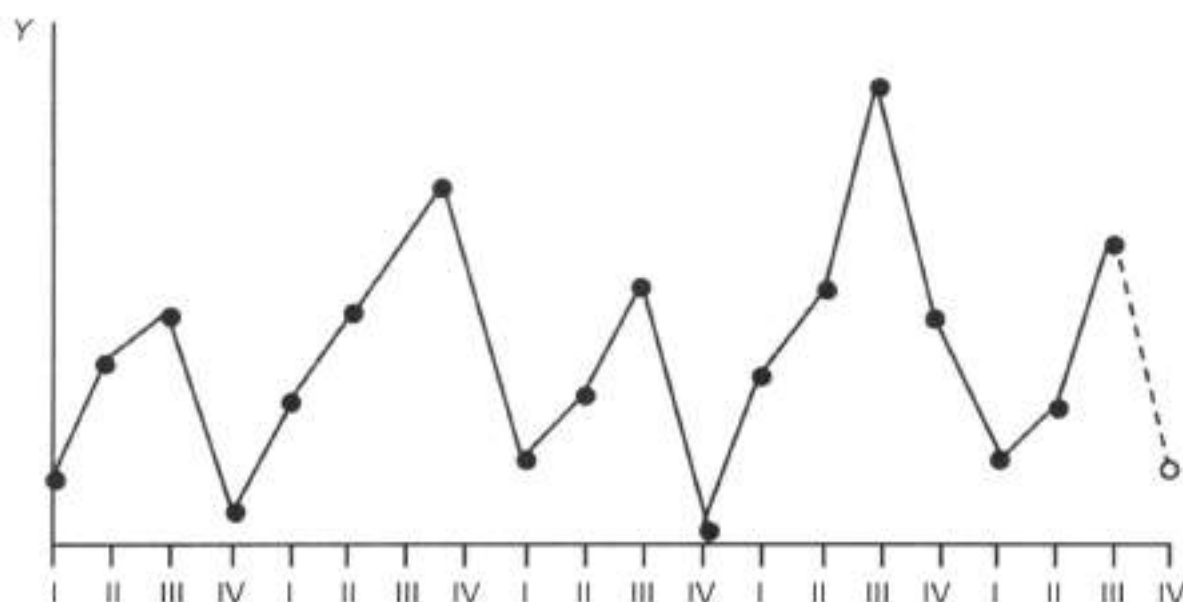
FIGURE 11.3 Horizontal, or No-Trend, Pattern

FIGURE 11.4 Seasonal Trend Pattern

A **cyclical trend** exists when long-term economic fluctuations associated with the business cycle influence the data. Thus, during a period of slow economic growth, housing construction will decrease. Most economic data tend to be sensitive to the business cycle and fluctuate according to the phase of the cycle. Figure 11.5 reflects the cyclical variability of a series.

Returning to our example of quarterly tourist data, we can plot the data again according to the effects of various elements in a series (see Figure 11.6). Now the patterns are quite clear. The series reflects an overall positive trend over time, although it shows strong seasonal fluctuations, along with effects arising from broader economic

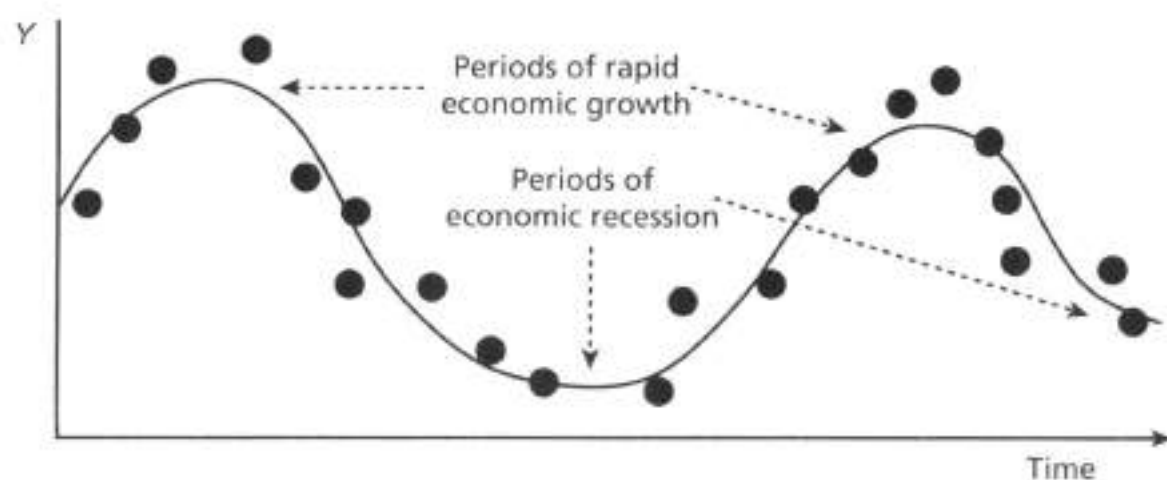
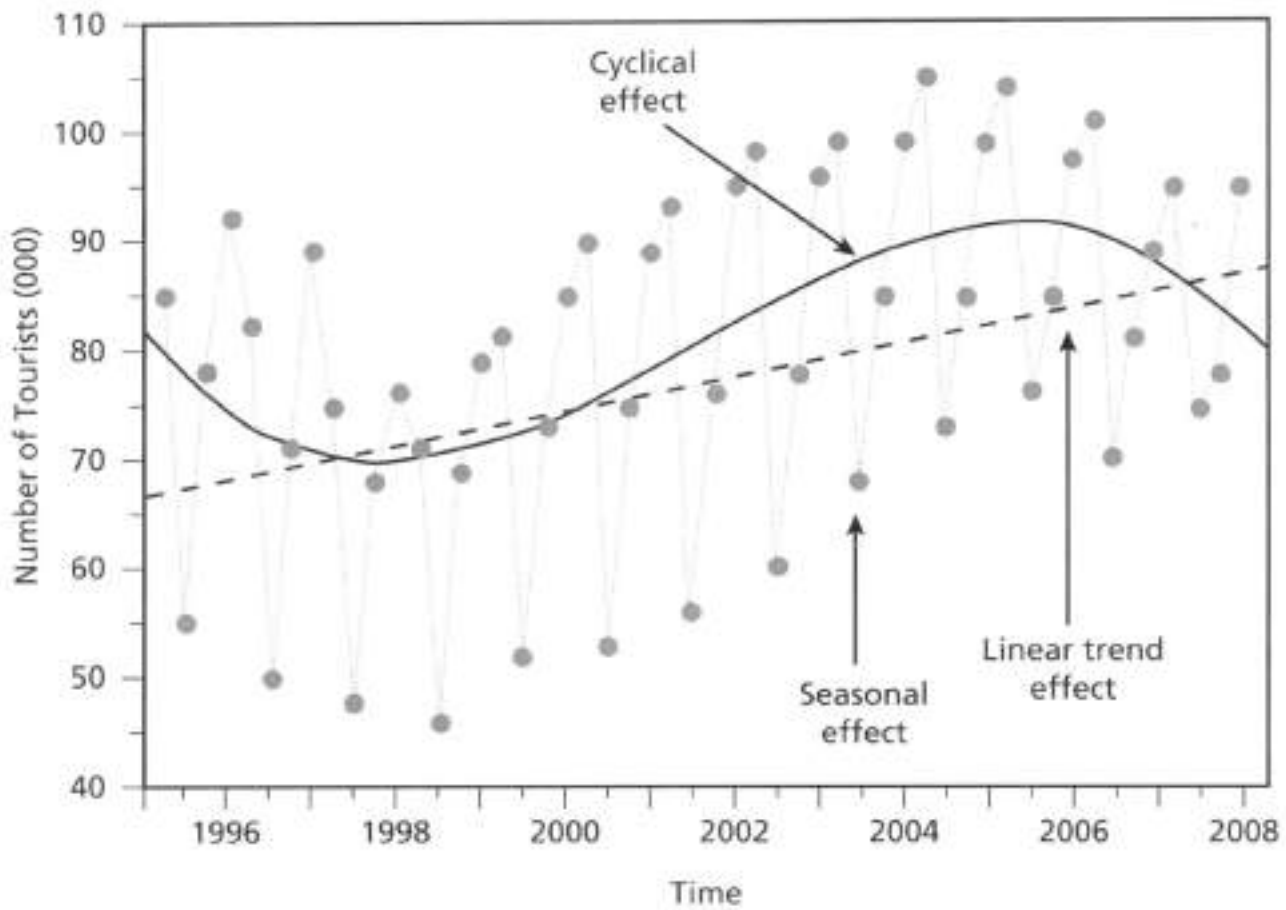
FIGURE 11.5 Effects of Economic Cycles

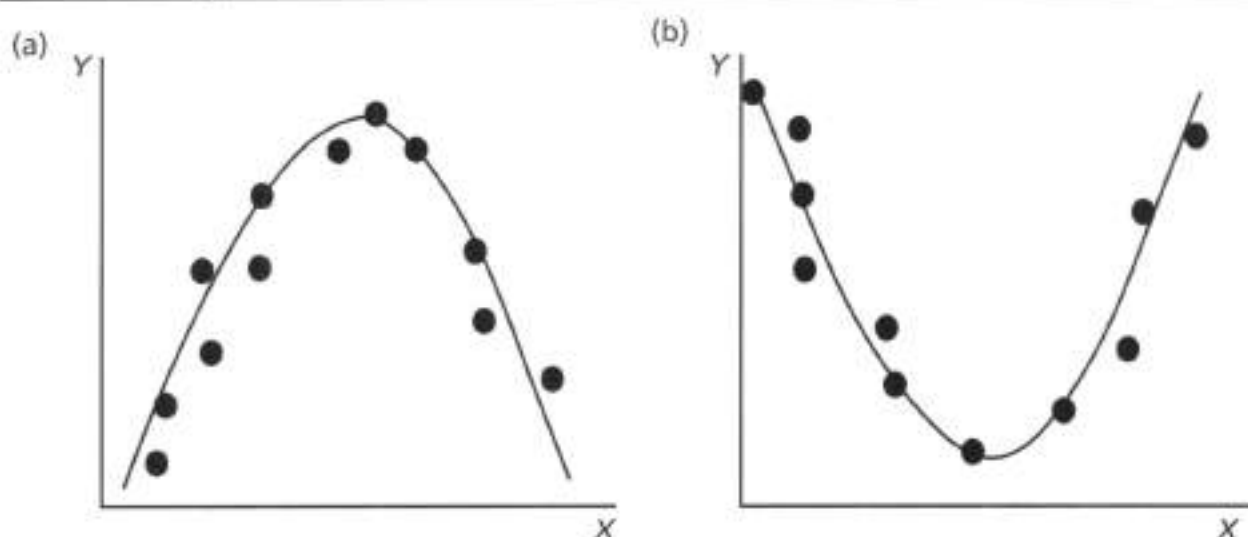
FIGURE 11.6**Discerning Trend Patterns in the Tourist Data from Masters, Pennsylvania**

trends over the thirteen-year period. Now that we have succeeded in showing definite patterns, our task of forecasting becomes a lot more manageable.

THE PATTERNS OF TIME TRENDS

Seasonal variations, the effects of a business cycle, and other external factors influence a data series. However, data may include a complex relationship with time that can cause a series to change directions in many different ways. Figures 11.7 (a) and 11.7 (b) show patterns of **quadratic trends**. The reverse U-shaped pattern is typical of a situation in which the dependent variable shows a pattern of increase followed by a decline. Data on the percentage of Americans living below the official poverty line demonstrate this kind of quadratic structure during the period from 1950 to 2000. The relation between agricultural production and the application of fertilizer is an example of a causal linkage in a quadratic trend. As one increases the use of fertilizers, the level of production tends to go up. However, after the point of saturation, these chemicals reach a toxic level and, consequently, begin to have a deleterious effect on production.

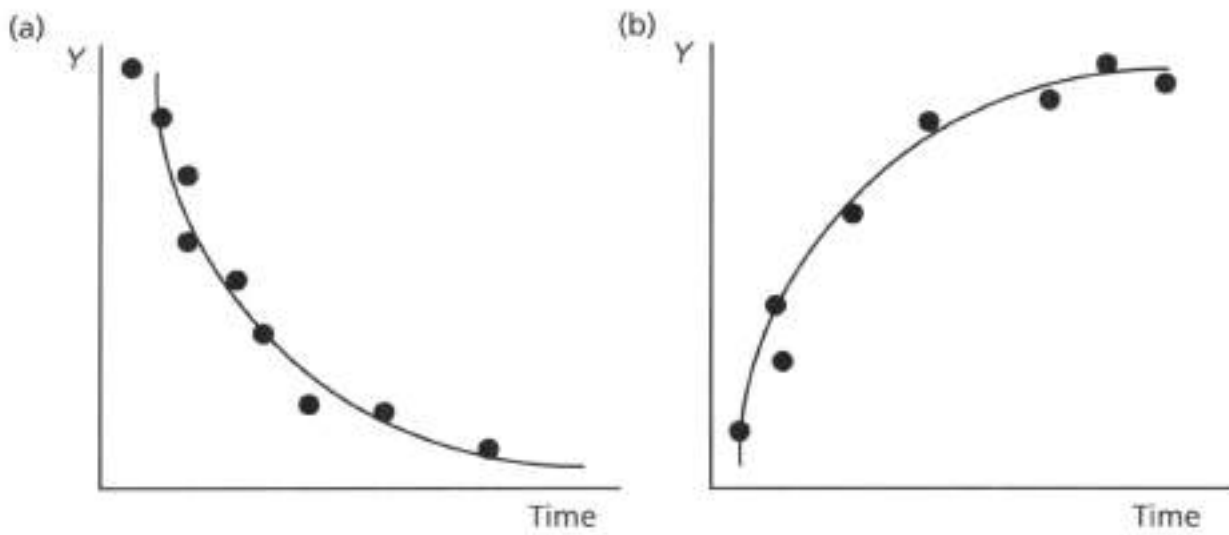
This U-shaped pattern is also typical of a per unit cost curve in a situation of increasing inefficiency as the scale of operation increases. Consider the operation of a

FIGURE 11.7 Quadratic Trends

county sheriff's department. If the sheriff's department operates on a very small scale, its cost of operation will be extremely high, because it will have to bear a large fixed cost for administration and other necessary operations. However, as the size of its operation increases (maybe because other small, incorporated cities start contracting their police work out to the sheriff's department), its per unit cost of operation (whichever way it is measured) will go down. But this downward trend cannot continue forever, and at a certain stage the increased size of the department will hamper its efficiency. Beyond this point (which in economics is regarded as the point of maximum efficiency), the unit cost of operation will start to increase. Therefore, by measuring the efficiency of an expanding organization over time, one may discern such a quadratic relationship.

A series may also depict an **exponential trend**, as shown in Figures 11.8 (a) and 11.8 (b). For example, viewed over time, population growth exhibits a negative exponential rate with income. That is, birthrates are highest among the underdeveloped nations in the world. However, these rates drop and then reach a plateau when nations attain high levels of economic prosperity. This pattern is shown in Figure 11.8 (a). In contrast, data on life expectancy show a positive exponential pattern similar to Figure 11.8 (b). Advancements in medical technology have increased people's life expectancy at a fairly rapid rate, but as we approach a biological limit on how long we can survive, the rate of growth slows down considerably.

A series can also depict a rather complex pattern of growth. For example, a series can have a **logistic trend**, or an S-shaped trend, such as the one shown in Figure 11.9. This trend shows a changing pattern of growth. The rate of growth is slow at the beginning of the series but changes during a transition period, at which time the rate of growth becomes quite rapid. After a certain point, this rate slows as the series reaches a steady state or an upper asymptote (a ceiling). This behavior is considered typical of a learning situation. When we try to learn something new (say, a foreign language), at the beginning our progress is painfully slow. After we have mastered the basics, our ability

FIGURE 11.8 Exponential Trends

to absorb material accelerates, only to reach a point of saturation when we attain the upper limits of learning.

A series is said to have demonstrated a **catastrophic trend** when there is a sharp discontinuity (see Figure 11.10). This kind of a precipitous rise or fall occurs as a result

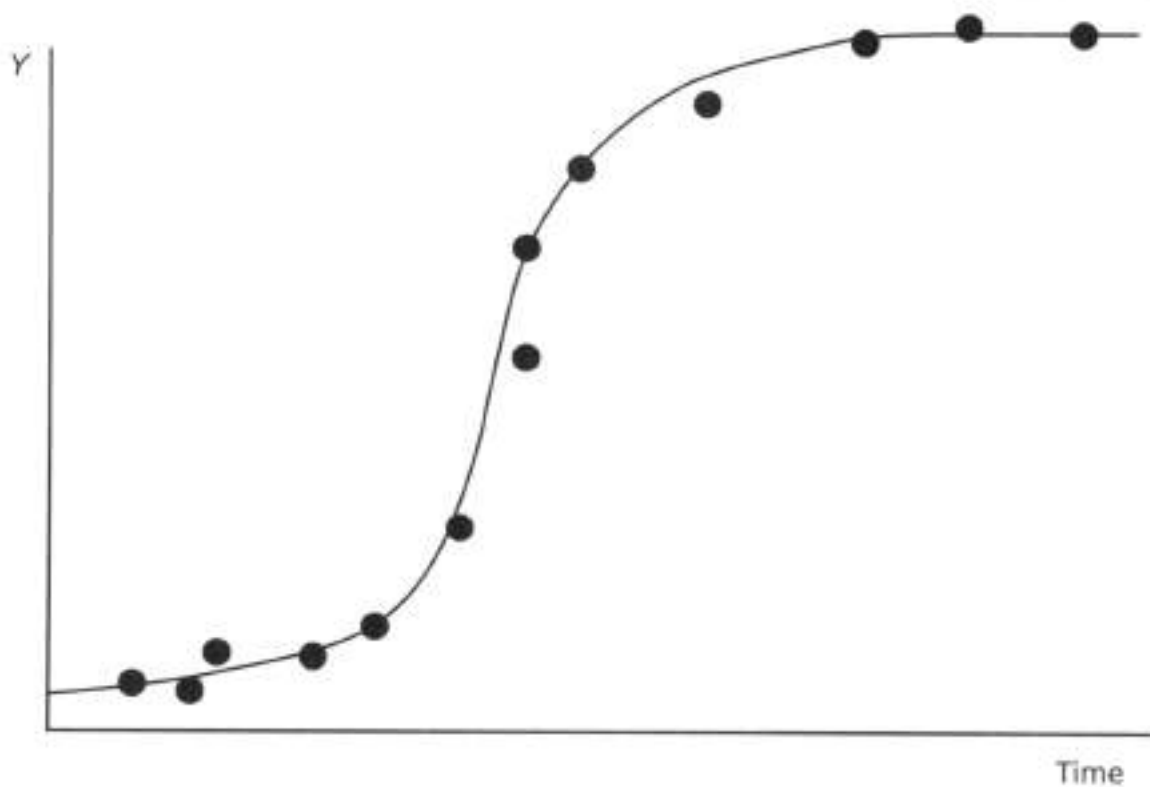
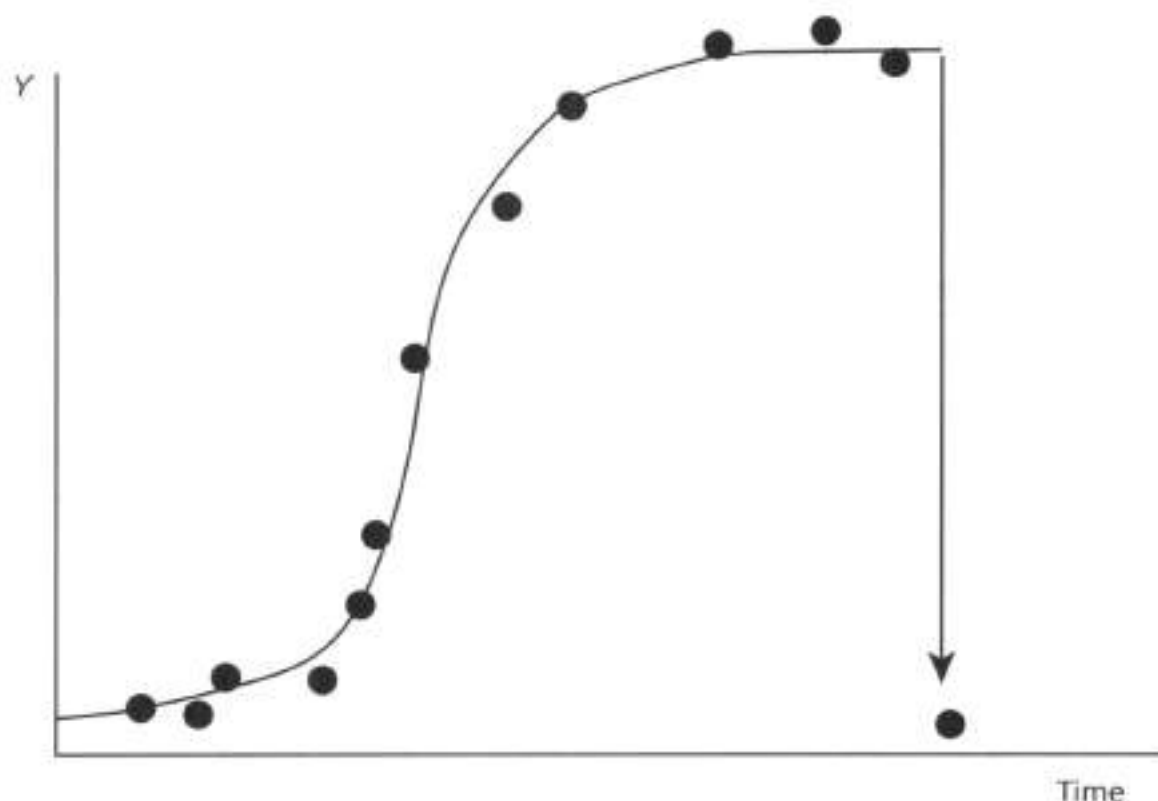
FIGURE 11.9 Logistic Trend

FIGURE 11.10 Catastrophic Trend

of a war or some other kind of national calamity. If one examines the growth of per capita GDP of the former Soviet republics, the catastrophic impact of the dissolution of the Soviet Union is apparent as their GDPs plunge. The technique of catastrophic **trend analysis** is still in its infancy and therefore is not commonly used in social science research or public policy analysis.

ADJUSTMENT METHODS

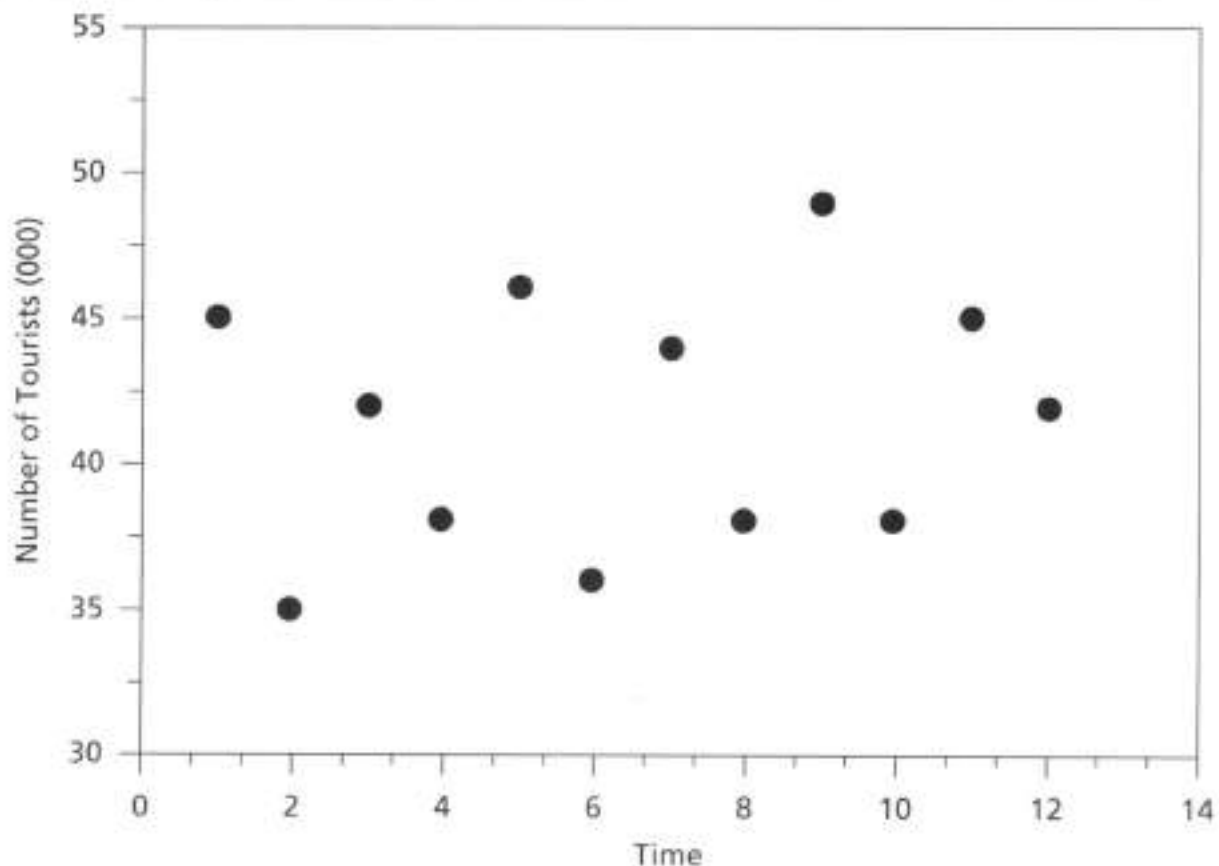
Looking again at Figure 11.1, we would probably agree that the true nature of a trend can be lost in the seasonal fluctuations of a series. In such cases, you may want to examine the data by filtering out the effects of these fluctuations. Suppose you want to know if the unemployment rate went up in December. The unemployment data will show the natural effects of a seasonal fluctuation because many types of outdoor work, from construction to agriculture, either stop or slow down during the winter months. To get a proper perspective, you will have to seasonally adjust monthly or quarterly unemployment figures. You may have noticed that the statement “seasonally adjusted” always qualifies quotations of unemployment statistics. Similarly, the presence of a strong trend can muddle the pure effects of seasonal fluctuations. That is, a rapidly growing economy can blunt the seasonal fluctuations in employment statistics. Unless this trend factor is eliminated, the effects of the seasons on the data may

TABLE 11.1**Tourist Population of Masters, Pennsylvania
(in thousands)**

Quarter	2006	2007	2008
I	45	46	49
II	35	36	38
III	42	44	45
IV	33	38	42

not be apparent. Perhaps the simplest way of eliminating the effects of seasonal fluctuations and trends is the method of seasonal and trend adjustment.

Let us return to Masters, our Pennsylvania resort town known for its winter and summer recreational facilities. The city's chamber of commerce is trying to understand the nature of tourist demand for its facilities. Consider the records of quarterly figures of tourist populations for the past three years (Table 11.1). The data presented in Table 11.1 do not reveal a great deal of information regarding the nature of the town's tourist industry. Neither does their plot, shown in Figure 11.11. Indeed, we would be hard-pressed to draw too many conclusions from the information we have at hand.

FIGURE 11.11**Plot of Number of Tourists per Quarter (in thousands)**

Seasonal Adjustment

The method of seasonal adjustment can lend a helping hand. The reason for this apparent confusion regarding the data is that two factors are at play: the seasonal effect and the trend effect. The presence of these two factors is obfuscating the picture presented by the raw data. Let us begin by suppressing the effects of seasonal variations to accentuate the effects of the trend. If you look at the data carefully, you will see that the first and the third quarters, which cover the peak skiing and summer seasons, tend to bring in more tourists. If there were no seasonal variation whatsoever, the quarterly totals would have been the same. Therefore, we can adjust for this variation by forcing each season to be the same (see Table 11.2).

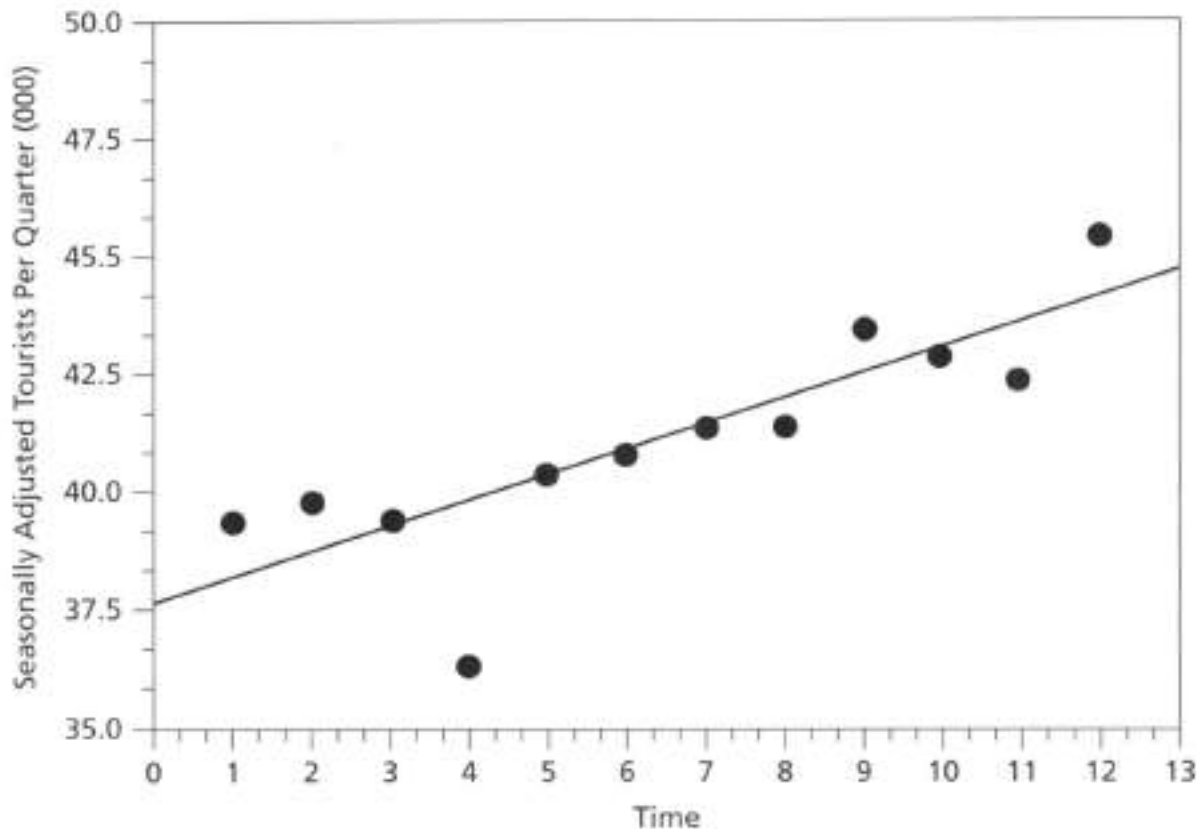
TABLE 11.2 Calculation of Quarterly Adjustment Factors (in thousands)

Quarter	Quarterly total (1998–2000)	Quarterly average	Quarterly adjustment factor (quarterly average minus grand average)
I	140	46.7	+5.6
II	109	36.3	–4.8
III	131	43.7	+2.6
IV	113	37.7	–3.4
Total	493	41.1	

The quarterly totals in the table were calculated by adding the number of tourists for each of the four quarters during the three-year period. During the first quarter, 45,000 tourists visited the town in 2006, 46,000 in 2007, and 49,000 in 2008. Therefore, during the first quarters of these three years, the town had a total of 140,000 visitors. By dividing the quarterly total by three, we calculated the average quarterly figure for the study period. You may notice that during this period there was a total of 493,000 visitors, or about 41,100 tourists per quarter. This is the average number of tourists we could expect if there were no seasonal effect. Therefore, we can calculate the **seasonal adjustment factor** by subtracting the grand average from each quarterly average. This is the adjustment factor, which, subtracted from each individual quarter, will give us a seasonally adjusted figure (see Table 11.3).

TABLE 11.3 Seasonally Adjusted Quarterly Tourist Population (in thousands)

Quarter	1998			1999			2000		
	No.	Adjustment factor	Adjusted no.	No.	Adjustment factor	Adjusted no.	No.	Adjustment factor	Adjusted no.
I	45	–5.6	39.4	46	–5.6	40.4	49	–5.6	43.4
II	35	+4.8	39.8	36	+4.8	40.8	38	+4.8	42.8
III	42	–2.6	39.4	44	–2.6	41.4	45	–2.6	42.4
IV	33	+3.4	36.4	38	+3.4	41.4	42	+3.4	45.4

FIGURE 11.12 Plot of Seasonally Adjusted Quarterly Data

Armed with this new series of numbers, the chamber of commerce can plan more effectively for the future, knowing that at least for the past three years, there has been a definite upward trend in the number of visitors coming to town. By comparing Figure 11.11 with Figure 11.12, you can see how this process of seasonal adjustment has subdued the quarterly variations and emphasized the trend pattern.

Trend Adjustment

The chamber of commerce may also want to understand the patterns of seasonal fluctuations without the distorting effects of a trend. Trends can be eliminated simply by adjusting the data for yearly variations. Again, the logic is the same. If there were no trend pattern, there would be no variation across time. In this situation seasonal factors will be the sole cause of fluctuations. We can calculate the yearly totals for the three years and then subtract each year's total from the yearly average to get the yearly adjustment factors, or the **trend adjustment factors** (see Table 11.4). These yearly adjustment factors show the average change from the mean. By adjusting the data for each quarter by this factor, we will be able to suppress the influence of the trend.

Now, by adjusting the yearly data, we can eliminate the effects of the trend factor and accentuate seasonal variations. This adjustment process can be seen in Table 11.5.

In Figure 11.13 the emphasis is on seasonal changes, and the trend factor has been filtered out. This diagram will help the policymakers in Masters visualize the seasonality of the tourist trend and plan for the future. Although the methods of seasonal and trend adjustment are not methods of forecasting per se, they can shed a good deal of light on

TABLE 11.4 Calculation of Yearly Adjustment Factors (in thousands)

Year	Yearly total	Yearly average per quarter	Yearly adjustment factor (yearly average minus grand average)
2006	155	38.8	-2.3
2007	164	41.0	-0.1
2008	174	43.5	+2.4
Total	493	41.1	

the behavior of a seemingly chaotic data series. As you can see in Figure 11.13, the trend-adjusted data show that the number of tourists visiting Masters in the first two quarters of the three years remained stable but that the number went up steadily for the fourth quarter. This insight into the data can be a powerful planning and analysis tool for developing appropriate public policies to bring more tourists to the area during the off-season. On the basis of this analysis, the city may hire a public relations firm to publicize the city's attractiveness to potential visitors for the first half of the year.

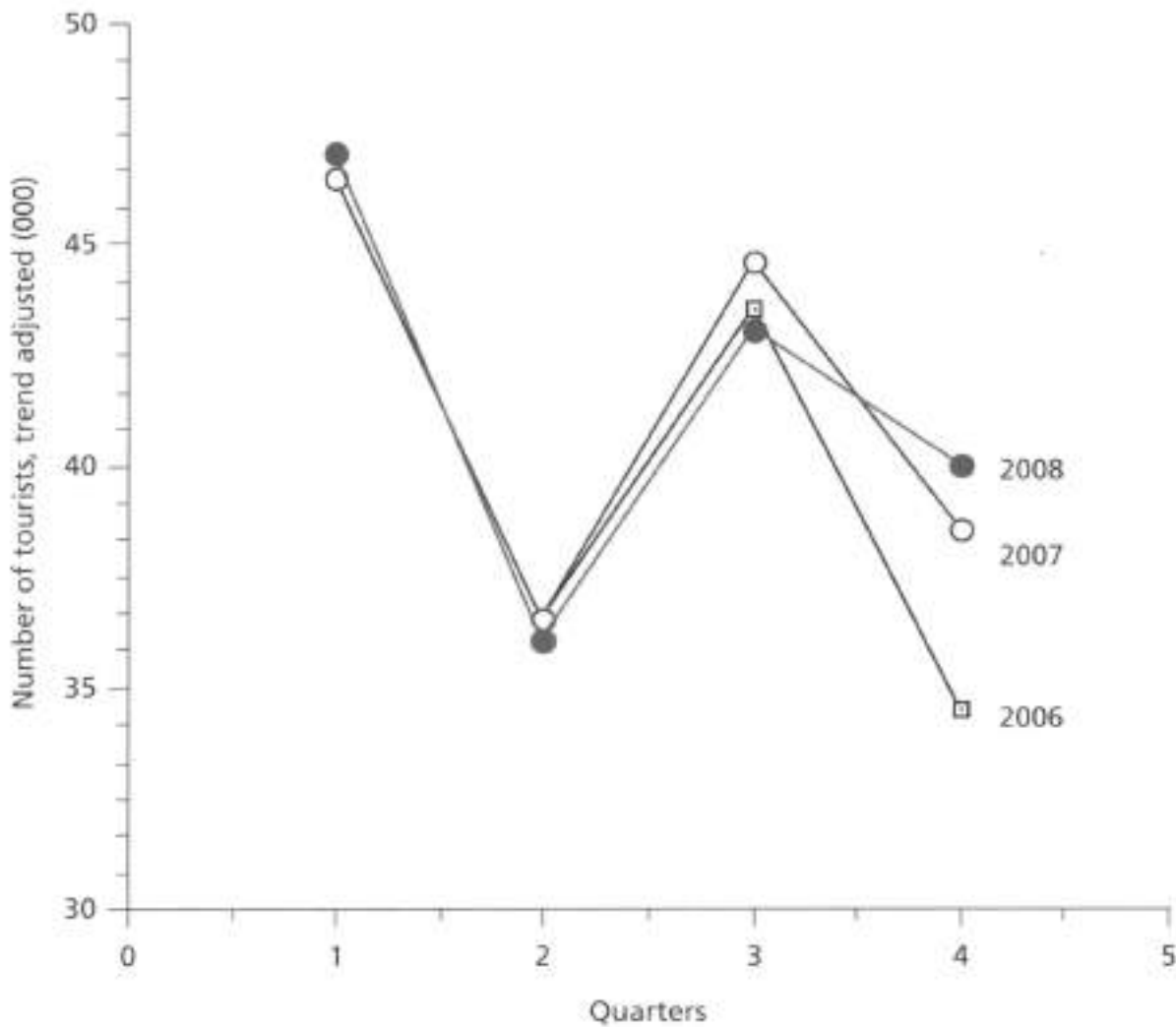
TABLE 11.5 Trend-adjusted Quarterly Tourist Population (in thousands)

Quarter	2006			2007			2008		
	No.	Adjustment factor	Adjusted no.	No.	Adjustment factor	Adjusted no.	No.	Adjustment factor	Adjusted no.
I	45	+2.3	47.3	46	+0.1	46.1	49	-2.4	46.6
II	35	+2.3	37.3	36	+0.1	36.1	38	-2.4	35.6
III	42	+2.3	44.3	44	+0.1	44.1	45	-2.4	42.6
IV	33	+2.3	35.3	38	+0.1	38.1	42	-2.4	39.6

SMOOTHING OUT THE FLUCTUATIONS

In the previous section, we discussed ways to smooth out seasonal fluctuations to reveal the underlying trend. However, when we want to consider a time span that is longer (say, a year or six months) or shorter (a week or a month, for example) than a season, or when fluctuations are more random than seasonal effects, we seek the help of a different method of smoothing. In these cases we use such techniques as **naïve projection** and **moving average**.²

Let us return to the quarterly time-series data on tourists visiting Masters during the thirteen-year period from 1996 to 2008 (see Table 11.6). I plotted this data series in Figure 11.1. As noted earlier in the chapter, the trend line is hidden behind fluctuations caused by seasonal variations and cyclical fluctuation. By examining this figure you can now clearly see a positive trend, the fluctuations caused by the four seasons, as well as the impact of a strong business cycle. Now we want to use this data series to forecast values for the next quarter. If we do not want to be restricted to the four quarterly figures, we will have to use one of the following methods of projection.

FIGURE 11.13 Plot of Trend-adjusted Tourist Data

Projecting the Immediate Past: Naive Projection

If I were to ask you to predict your earnings for the coming year, your prediction would probably be based on what you earned this year. Prediction based on the last period's performance is often called the naive method of projection. In Table 11.6, the quarterly data are listed in the first column. In statistics this process is called **lagging**. If we lag a column by one period, we express it in terms of what happened in the immediately preceding period. Similarly, a three-period lag would present the series in terms of what happened three periods ago. When we write in algebraic symbols, we express periods in subscripts. For example, X_t would mean the value of the variable X in period t , which in this case is the current year. Similarly, X_{t-1} would mean last period's value (or the value lagged by one period), and X_{t-10} would be the value lagged by ten periods.

Because an organization's decisions are almost always made on the basis of incremental reasoning (an argument such as, "Last year this program cost \$150,000; therefore, this year we expect it to cost . . ."), forecasting on the basis of what happened last period is probably the most frequently used method. This method produces a reasonably good forecast if there is little trend, positive or negative, and there are few fluctuations. If you have steady salaried employment, your current year's earnings would serve as a good proxy for the next year's prediction. If your income tends to fluctuate, however, this

TABLE 11.6**Quarterly Series on the Number of Tourists Visiting Masters, Pennsylvania**

Year	Quarterly data	Lag (1 year)	Moving average (2 years)	Moving average (4 years)
1996.1	85	—	—	—
1996.2	55	85	—	—
1996.3	78	55	70.0	—
1996.4	92	78	66.5	—
1997.1	82	92	85.0	77.50
1997.2	50	82	87.0	76.75
1997.3	71	50	66.0	75.50
1997.4	89	71	60.5	73.75
1998.1	75	89	80.0	73.00
1998.2	48	75	82.0	71.25
1998.3	68	48	61.5	70.75
1998.4	76	68	58.0	70.00
1999.1	71	76	72.0	66.75
1999.2	46	71	73.5	65.75
1999.3	69	46	58.5	65.25
1999.4	79	69	57.5	65.50
2000.1	81	79	74.0	66.25
2000.2	52	81	80.0	68.75
2000.3	73	52	66.5	70.25
2000.4	85	73	62.5	71.25
2001.1	90	85	79.0	72.75
2001.2	53	90	87.5	75.00
2001.3	75	53	71.5	75.25
2001.4	89	75	64.0	75.75
2002.1	93	89	82.0	76.75
2002.2	56	93	91.0	77.50
2002.3	76	56	74.5	78.25
2002.4	95	76	66.0	78.50
2003.1	98	95	85.5	80.00
2003.2	60	98	96.5	81.25
2003.3	78	60	79.0	82.25
2003.4	96	78	69.0	82.75
2004.1	99	96	87.0	83.00
2004.2	68	99	97.5	83.25
2004.3	85	68	83.5	85.25
2004.4	99	85	76.5	87.00
2005.1	105	99	92.0	87.75
2005.2	73	105	102.0	89.25
2005.3	85	73	89.0	90.50
2005.4	99	85	79.0	90.50
2006.1	104	99	92.0	90.50
2006.2	76	104	101.5	90.25
2006.3	85	76	90.0	91.00
2006.4	97	85	80.5	91.00
2007.1	101	97	91.0	90.50
2007.2	70	101	99.0	89.75
2007.3	81	70	85.5	88.25
2007.4	89	81	75.5	87.25
2008.1	95	89	85.0	85.25
2008.2	75	95	92.0	83.75
2008.3	78	75	85.0	85.00
2008.4	95	78	76.5	84.25
2009.1 (forecast)	—	95	86.5	85.75

method will not give you a very good forecast. For example, because seasonality causes ample variations, the lagged data presented in Table 11.6 act as a poor predictor for the following quarter.

Projecting by the Mean

Another method of projection is calculating the mean. For instance, consider a series:

$$X_1, X_2, X_3, X_4, \dots, X_{t-1}, X_t$$

In this case we can make a prediction for the following period, X_{t+1} , simply by taking the mean of the series—that is, by calculating

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_{t-1} + X_t}{t} = \hat{X}_{t+1} \quad (11.1)$$

where t is the total number of observations and $\hat{X}_{t+1}$ is the projected value for the period $t+1$.

Therefore, the forecast of X_{t+1} is given by the mean of the series

$$\hat{X}_{t+2} = \frac{X_1 + X_2 + \dots + X_t + X_{t+1}}{t+1}$$

You might wonder when this simple method is appropriate for forecasting a series. This method is useful only if you have a stationary series with no observable trend (see Figure 11.3). In such cases, the mean of the series would give you the best possible forecast for future behavior. When predicting fluctuations in rainfall, you may do best by taking the yearly average. The same may be true for predicting price movements for a specific period in the stock market.

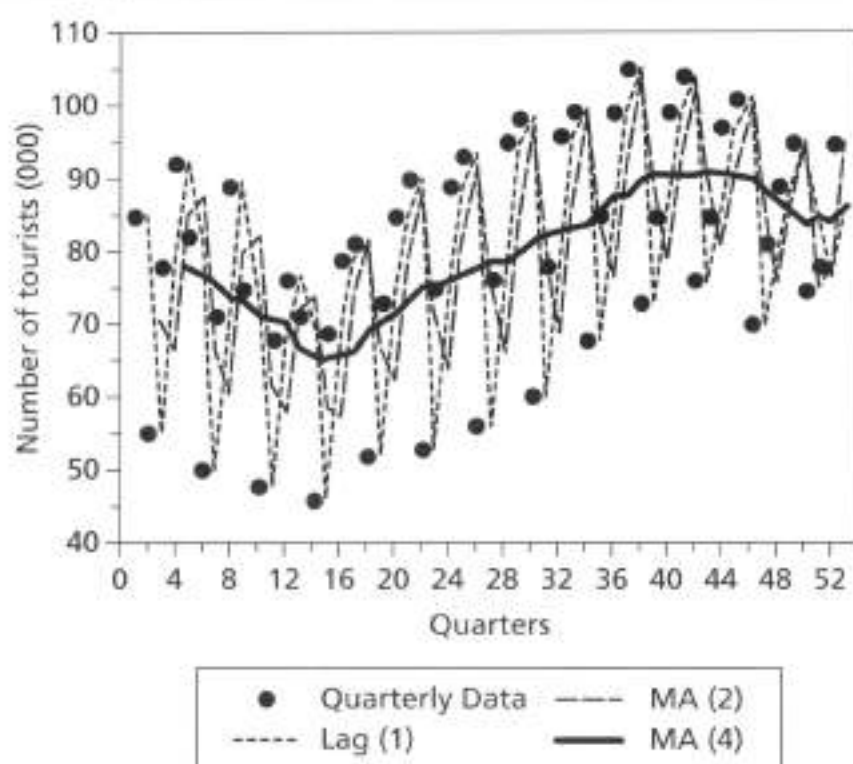
Moving Average

When there is a clear trend, the variations within a series can be ironed out by what is known as the method of moving average. A moving average is calculated by averaging two or more consecutive values in the series and accepting the computed value to be the forecast for the next period, as shown in Table 11.6. Using this method, we would predict for X_{t+1} as follows:

$$\hat{X}_{t+1} = \frac{X_t + X_{t-1}}{2} \quad (11.2)$$

Lagging the data by one period is called the moving average of order 1, or the first-order moving average. As you can see in Table 11.6, when calculating the moving average of order 2, we lost information on the first two observations because these were used to calculate the forecast for the third period. We can increase the order by taking the average of a larger number of lagged data. If we want to calculate a fourth-order moving average, $\hat{X}_{t+1}$ will be computed as

$$\hat{X}_{t+1} = \frac{X_t + X_{t-1} + X_{t-2} + X_{t-3}}{4} \quad (11.3)$$

FIGURE 11.14**Adjusted and Unadjusted Data on Tourists Visiting Masters, Pennsylvania**

Note: MA = moving average.

If you have quarterly data that reflect a strong seasonal effect, a four-period moving average would essentially iron out the seasonal fluctuations. This is shown in Figure 11.14. As the order of moving average increases, the top of the series loses more information, and the series shows a greater degree of smoothing out.

The moving average method requires you to decide on the number of periods to use to smooth the series. You can see from Figure 11.14 that if you want your newly created series to look like the original data, you should choose a smaller lag period. Hence, if a series has a clear trend with a relatively small degree of fluctuation, use a short (perhaps a two-period) lag. However, if the series contains a large degree of seemingly random or seasonal variations, use a larger lag period. Although the choice is highly subjective and dependent on the needs of your client, for the series under discussion, given the high degree of seasonal variation, I would choose a longer lag, perhaps a four-period one.

When you do a trend analysis, it is imperative that you proceed in a systematic way. Here is a quick list of steps you can follow when conducting your analysis:

1. Specify the objective of your study.
2. Plot the data.
3. Look for a trend pattern.
4. If the series is monthly or quarterly, look for seasonal variability.
5. For a long yearly series, look for cyclical effects.
6. Correct for seasonality or trend, depending on the purpose of the study.
7. If there is no trend, use the mean for forecasting.

Choice of Projection Technique

In this chapter we have discussed several different projection techniques. As you can see, each method has its pros and cons. Note that you may obtain a different forecasted value depending on the technique you choose to use. Although there is no definitive roadmap for selecting the “best” method, I suggest following these steps:

1. Define the objective of projection.
2. Plot the data.
3. If there is a great deal of seasonal variability, for comparison purposes, do seasonal and trend adjustments.
4. If there is no trend, use the mean of the series.
5. If there are seasonal fluctuations, use the moving average. Remember, the higher the order of the moving average, the greater the smoothing effect.

THE POLITICS OF FORECASTING

If forecasting on its own merits is a tricky business, forecasting for public policy can be even more problematic. Because a forecast may involve the livelihoods of many individuals and has the potential to cause huge monetary damages, a forecast should be of great concern. Take, for example, the case of Mammoth, a mountain resort town in the picturesque eastern Sierra region in California. Like many other places in the state, this region is susceptible to earthquakes. To complicate the matter, Mammoth Mountain is a dormant volcano. In 1998 the Hollywood movie *Dante's Peak*, set in a mythical resort town meant to mimic Mammoth, showed the plight of the town when the volcano suddenly erupted. This sensational movie, along with some dire predictions from volcanologists, caused the ski town a great deal of anxiety as real estate prices threatened to fall faster than the skiers on its slopes. David Hill, chief scientist with the U.S. Geological Survey and a longtime observer of the region, calmed frayed nerves with his prediction that the odds of a sudden volcanic eruption were grossly overstated.³

Although forecasts are supposed to be accurate descriptions of the future, the power they hold in the area of public policy can often make them tools of special interest groups. The current debate on forecasts relating to global warming provides an example of the issues involved. If the cause of global warming could be pinned to the emission of hydrocarbons and other greenhouse gases resulting from human activity, public policies would severely crack down on these polluting agents. Since such agents, in the form of large corporations, are concentrated in the industrial West, such a forecast can be highly significant for its economy.

Forecasts can also be used by scientists and environmentalists for the purpose of raising awareness. In the late 1950s a group of environmentalists became engaged in studying the entire world as one single socioeconomic and environmental system.⁴ Their work alerted a generation of young scientists and environmental activists to the dangers of relying on dwindling natural resources. Although the claims of a catastrophe were clearly overblown (and were realized as such at the time of the study), their research was instrumental in educating people about the risks of unrestrained growth. (For a report on how forecasting has affected the debate over the budget surplus, see “A Case in Point.”)



A CASE IN POINT

Fighting over a Fiction? The Budget Surplus in the Presidential Debate

During the fierce 2000 presidential campaign, one issue dominated the national debate: how to spend the \$4.6 trillion national surplus projected over the next ten years.⁵ On the basis of this rosy prediction, Republican George W. Bush wanted to give a third of the surplus back to the taxpayers and use the rest on national defense and education, as well as pay down the national debt. Democrat Al Gore wanted to pay down the national debt, as well as spend on numerous other social and military programs. Gore promised to eliminate the federal debt by 2012; Bush claimed he would do the same by 2016.

While these candidates exchanged barbs and insults and promised the most judicious spending plans, some experts quietly questioned the veracity of the forecasts. Because the budget surplus was based on the performance of a hot economy in 1999, experts feared that if the economy did not keep up with exceptional years of growth, the projected surplus would quickly melt away. In fact, if the economy produced a lower surplus, or no surplus at all, any ambitious spending plan would take the economy back to the days of deficit spending and the consequent burgeoning public debt.

To make the situation worse, the Congressional Budget Office warned that after the next ten years, the government would face a funding crunch from Social Security, Medicare, and Medicaid that could "drive federal debt to unsustainable levels." As the aging baby boomers would demand increasingly more health care and other social service dollars, around 2020 the federal budget

KEY WORDS

Catastrophic trend (p. 255)
 Cyclical trend (p. 252)
 Exponential trend (p. 254)
 Horizontal trend (p. 251)
 Lagging (p. 261)
 Linear trend (p. 250)
 Logistic trend (p. 254)
 Moving average (p. 260)
 Naive projection (p. 260)
 Negative trend (p. 249)

Positive trend (p. 249)
 Projecting by the mean (p. 263)
 Quadratic trends (p. 253)
 Random errors (p. 250)
 Seasonal adjustment factor (p. 258)
 Seasonal trend (p. 251)
 Stationary series (p. 251)
 Trend adjustment factors (p. 259)
 Trend analysis (p. 256)

would be overwhelmed by these limitless demands. And it could get worse. The projected surplus was divided into two parts: \$2.4 trillion in the Social Security program and \$2.2 trillion in the rest of the budget. The analysts were comfortable with the projection for Social Security because it was based on relatively stable demographic data. However, it was the surplus in the rest of the federal budget that worried independent observers. On the basis of a different set of assumptions, the Center on Budget and Policy Priorities, a liberal group, estimated that the ten-year surplus outside Social Security would come closer to \$700 billion. The Brookings Institution, a somewhat more centrist think tank, estimated that the figure might be as low as \$352 billion. If these lower projections came true, the budget surplus would not be much of an issue to fight over.

This "Case in Point" appeared in the first edition of this book, published in 2001. Today, nearly a decade later, the discussion of eliminating the federal public debt by 2012 or even 2016 seems a distant dream from a different land. At the time of writing the present edition, the current year budget deficit—resulting from an economic meltdown, massive bailouts of the banking and automobile industries, and two long wars in Iraq and Afghanistan—is reaching \$1 trillion for the first six months of fiscal year 2009, and the total public debt stands at more than \$11 trillion and counting.⁶

Discussion Points

1. Why do experts differ on forecasts?
2. What are the political pitfalls of public policy forecasting?
3. Can you ensure an "unbiased" forecast when it comes to public policy?

EXERCISES

1. Write an essay on trend analysis when the past behavior is known. Give examples from real life for the various kinds of trend patterns discussed in this chapter.
2. A typical time series contains wide-ranging fluctuations. Explain, with examples, various sources of fluctuations in a data series. How can you adjust for such fluctuations?
3. After reviewing the U.S. Government Accountability Office report on poverty in America in Appendix A, inspect the U.S. Census Bureau's data on poverty. Using the moving average, can you project the U.S. poverty rate from the Census Bureau's data?
4. In 1987 several states increased the speed limit on their highways in sparsely populated areas. A county analyst is trying to see if the change caused a discernible

increase in the accident fatality rate. She has compiled a table of quarterly fatality data:

Quarter	1985	1986	1987	1988
I	85	76	88	92
II	43	42	52	48
III	51	57	71	78
IV	98	97	110	121

What conclusions can you draw for policy prescription?

5. You are in charge of investing your agency's liquid cash. You are considering investing in a particular security portfolio that has yielded the monthly returns shown below. How would you predict its behavior for the next three months?

Month	Yield
1	7.0
2	3.7
3	1.8
4	0.5
5	9.5
6	0.1
7	2.3
8	6.3
9	3.2
10	0.9
11	4.1
12	9.3
13	7.3
14	7.8
15	6.6

6. The table below shows the number of auto thefts in the city of Masters for the past fifteen years. Use the moving average technique to discern the overall trend and predict the number of auto thefts for the coming year. In this context, explain your choice of lag period.

Year	Auto thefts
1	98
2	87
3	110
4	112
5	108
6	121
7	132
8	125
9	127
10	130
11	145
12	153
13	148
14	151
15	149

7. Collect a series of data on any event of national, state, or local importance. Write a report on its trend pattern and predict its value for the next period.

Notes

1. Also, in reality, a series can assume a much more complex trend pattern. For a more detailed discussion, see Spyros Makridakis, Steven C. Wheelwright, and Victor E. McGee, *Forecasting: Methods and Application*, 2nd ed. (New York: Wiley, 1983).
2. To look at the use of moving average in discerning the underlying trend (and to help you forecast the future trend), log on to the Web page of any Wall Street brokerage firm (for example, Charles Schwab, at www.schwab.com). By choosing the box for a particular company (say, IBM), you can readily see the use of moving average as an aid to the mental forecasting of stock prices.
3. See Robert Lee Holtz, "Southland's Quake Danger Forecast Is Cut," *Los Angeles Times*, March 18, 1998.
4. See Donella Meadows, Dennis L. Meadows, Jorgen Randers, and William W. Behrens III, *The Limits to Growth* (New York: Signet, 1960).
5. Ken Moritsugu and Jackie Koszczuk, "Surplus Could Be a Figment of Faulty Estimates," *San Diego Union-Tribune*, October 7, 2000.
6. Jeff Bater, "U.S. Budget Deficit for 2009 Nears \$1 Trillion," *Wall Street Journal*, April 10, 2009, <http://blogs.wsj.com/economics/2009/04/10/us-budget-deficit-for-2009-nears-1-trillion>.

12

Projection Techniques: The Methods of Simple and Multiple Least Squares

Forecasting is integral to the process of planning. In the previous chapter we discussed the problem of projecting a trend when a long history is unavailable. However, in the current information age, we frequently encounter a long series of past information. When such information is available, forecasting becomes much more systematic, or “scientific” if you will, than forecasting on the basis of a single rate of growth or on subjective judgment.

We do not have to look far to find examples of forecasting in the public sector. State and federal governments routinely forecast revenues from various sources for the preparation of budgets. Local governments, unless they are extremely large, depend primarily on state governments to provide them with such information and do not generally get involved in projecting tax revenues. However, local government agencies frequently engage in forecasting needs for their services. A financial management department may seek permission to raise money from a bond issue for enlarging children’s recreation facilities. This may call for a forecast of the number of children in the community for the next five years. In addition to government agencies, branches of the federal reserve banks, research institutes, information consultants, and many other groups get involved in the business of forecasting the future.

We discussed the difference between a trend projection and causal analysis in chapter 10. In this chapter, we will examine one of the most commonly used techniques of trend projection based on a series of data over time (called a time series): the **method of least squares**, or regression models.¹ In a straight or linear trend, we use time as a single independent variable. If you plot the per capita gross domestic product (GDP) of the United States, you will see that the series shows a fairly steady linear trend pattern. I have plotted the data in Figure 12.1. In cases in which there are linear trends, we use **simple regression** models, which have only one independent variable (in our example,

time). However, at times a series may exhibit a complex, nonlinear trend pattern. In such a case, you may have to use a nonlinear model with more than one independent variable. These are called **multiple regression** models. The focus of this chapter is on the effects of time. Hence, we will discuss multiple regression within the context of a more complex time pattern in the second half of the chapter. (See Appendix B for a discussion of multiple regression as a part of building causal models.)

THE LOGIC OF THE LEAST SQUARES METHOD

The task of forecasting becomes considerably easier when we have long time-series data that allow us to see if there is a definite trend. When the trend is readily apparent, we can build a model that will help us forecast the future. For example, we can hypothesize from Figure 12.1 that GDP per capita is growing in a linear trend pattern over time. In other words, the elapsed time since the beginning of the study period determines the level of per capita income. The straight line drawn through the yearly data approximates the trend line. By our model we hypothesize that the per capita GDP of the country will continue to grow along this line and, by extending it into the future, we can forecast the future levels of per capita GDP.

In building a model for the projection of a trend, we hypothesize that the fundamental direction in a data series can be explained by the functional relationship proposed by the model. I can best explain the basic idea of a regression model with the help of Figure 12.2. The model in this figure shows the dependent variable (in this case, per capita income) as an input being transformed by the internal logic (or the functional relationship) of the system. An input stimulus (a change in the independent variable—in this case, the passage of time) produces the output (the new level of per

FIGURE 12.1

U.S. Gross Domestic Product per Capita (in 1972 constant dollars)

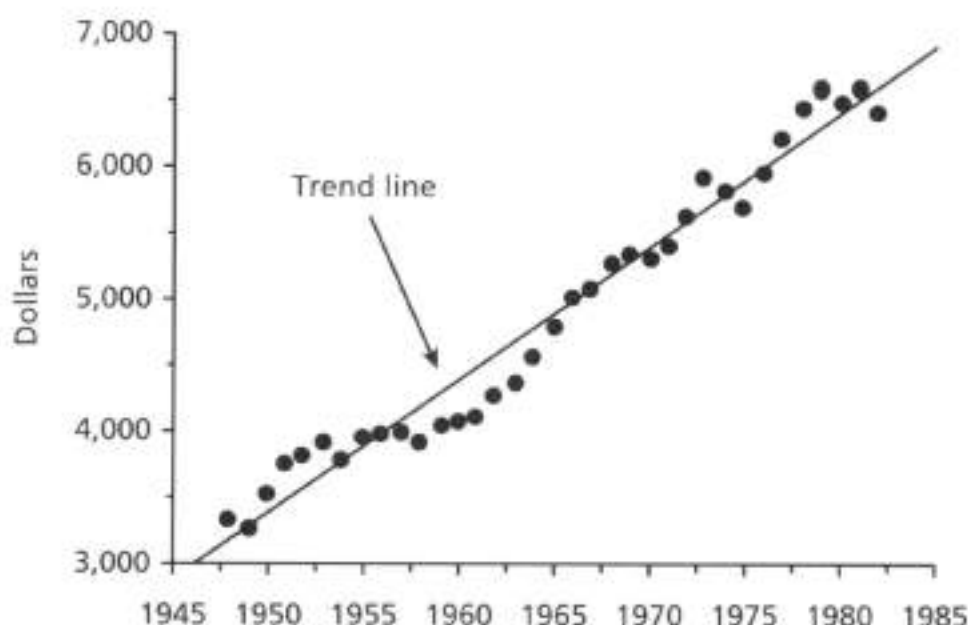
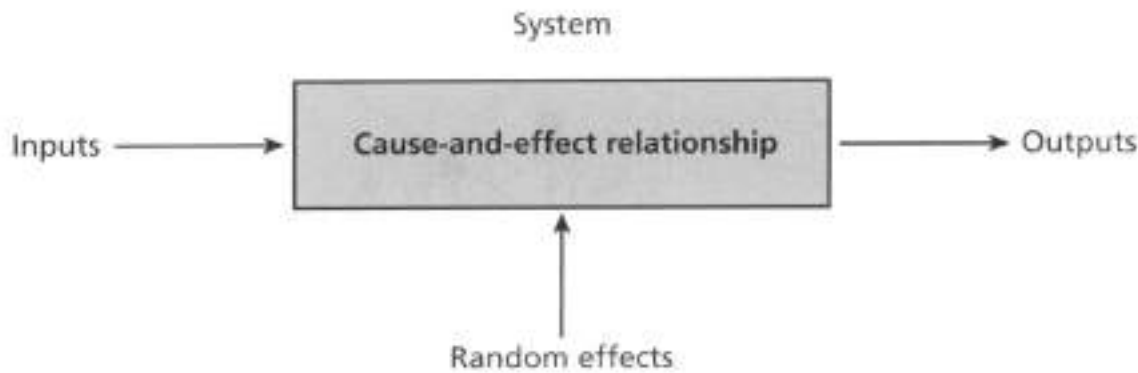


FIGURE 12.2 Time-series Model

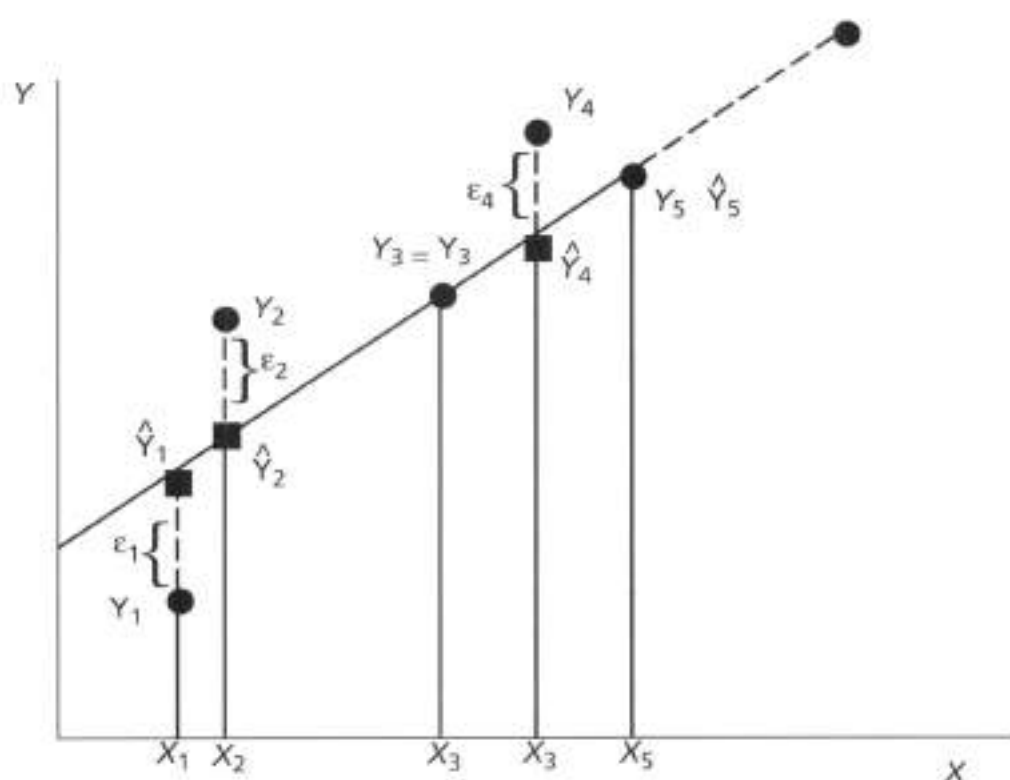
capita GDP). At this point, it is extremely important to remember that when we are including time as the independent variable, we are in fact claiming to be agnostic regarding the true causal relationship that is propelling the direction of the dependent variable. Thus, multiple factors could affect the growth of the per capita GDP of a nation, including world business cycles, the invention of new technologies, and new areas of trade and commerce (for example, in the case of the United States, trade with a more open China and the former communist nations). War, revolution, or cataclysmic natural disasters, such as earthquakes and hurricanes, are also likely to affect the growth of national domestic products. The basic assumption of a time-series model is that the simple passage of time embodies within it all these complex events, resulting in a steady discernible pattern, which we call a trend.

Despite our hypothesis regarding a simple linear trend, as you can clearly see in Figure 12.1, not all the actual observations fall on the trend line. If they did, we would have a perfect relationship, offering a perfect forecast of the future values. However, one rarely sees a perfect relationship in the realm of the social sciences. When some of the data points do not fall on this line, we claim that the disparity between our forecast and the actual observation is caused by **random errors**. These random events are defined as the effects of those independent variables that we did not include in our model.

Therefore, we assume that the observed system contains two components: the explained and the unexplained. That is,

$$\text{Data} = \text{Trend} + \text{Error}.$$

The primary purpose of trend analysis is to uncover the pattern that allows us to project it into the future. In this endeavor, we look for the “best line” that approximates the trend in a data series. In Figure 12.3, for example, we are measuring the values of the dependent variable Y on the vertical axis, while the horizontal axis measures the independent variable X . The plotted observations show a linear trend pattern. By extending the trend (shown by the dashes), we can project the future value of Y on the basis of X . Then the question becomes how to derive the trend line. It should be obvious that to obtain the line that best explains the trend, we must find the one that *lies closest to the*

FIGURE 12.3 Errors of Estimation

observations, or leaves us with the least amount of unexplained error. Errors are the differences between the projected trend line and the observed data points.

Suppose we have five observations on the dependent variable, Y , the trend of which we are trying to estimate by deriving a line as close as possible to the data points. Suppose we have obtained the trend line for the data series. The points on the trend line are the estimated values. For the first observation, Y_1 , we estimate it to be $\hat{Y}_1$ ("Y hat one"). Since our estimate falls short of the actual value, the difference between the actual and the estimated values is the random error, or the error of estimation, which we call ϵ_1 .

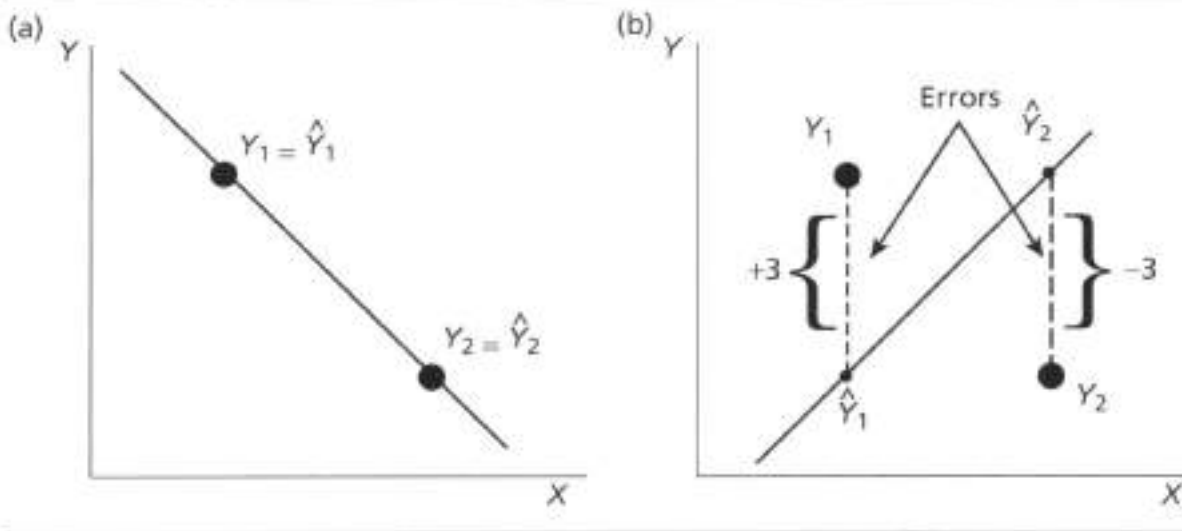
It is obvious that drawing the line of estimation by hand (as in Figure 12.3) will not do, because we can never be sure that it is in fact the closest line to the data. However, we can understand intuitively that the best possible line to fit a data set will have to be the least distance from the mean of the series. Consider the following data points:

$$X = 1, 2, 3, 4, 5$$

$$Y = 4, 6, 7, 9, 10$$

If you plot these points on graph paper, you can see that any line that does not go through the point $(3, 7.2)$ —3 being the average of the values of X and 7.2 being the average of the values of Y —cannot be considered closest to the data points.

However, our problem does not go away when we draw a line through the mean values of the two series. Since the deviation from the mean is always the least, the sum total of deviation from the mean is always equal to 0. Therefore, in some cases this rule

FIGURE 12.4 The Problems of Summing the Errors

will not allow us to identify the “best” line. This problem can be shown with the help of Figure 12.4. Suppose I have only two observations, and I want to draw the line to lie closest to these two points of observation. Clearly, as shown in Figure 12.4 (a), the line drawn through the two points will be the best line because for each observation, the error terms are equal to 0. However, suppose I draw this line in the opposite direction, as shown in Figure 12.4 (b). Let us assume that the distances from the trend line to the observed points are 13 and -3 , respectively. The problem with simply summing the errors is that the positive errors cancel out the negatives. If we want to choose the best line on the basis of the fewest number of errors, we are at a loss to choose between these two lines, because the sum of errors for both of them is equal to 0. How do we choose between them?

In chapter 6 we discussed the problem of deviations from the mean being equal to 0. Recalling our discussion, we can see that this problem can be avoided if we *square the deviations* from the line. By following this method, the line in Figure 12.4 (a) still gives us an error value of 0, whereas the sum of the squared deviations for the second line is $(+3)^2 + (-3)^2 = 18$. Clearly, we can now choose the first line over the second.

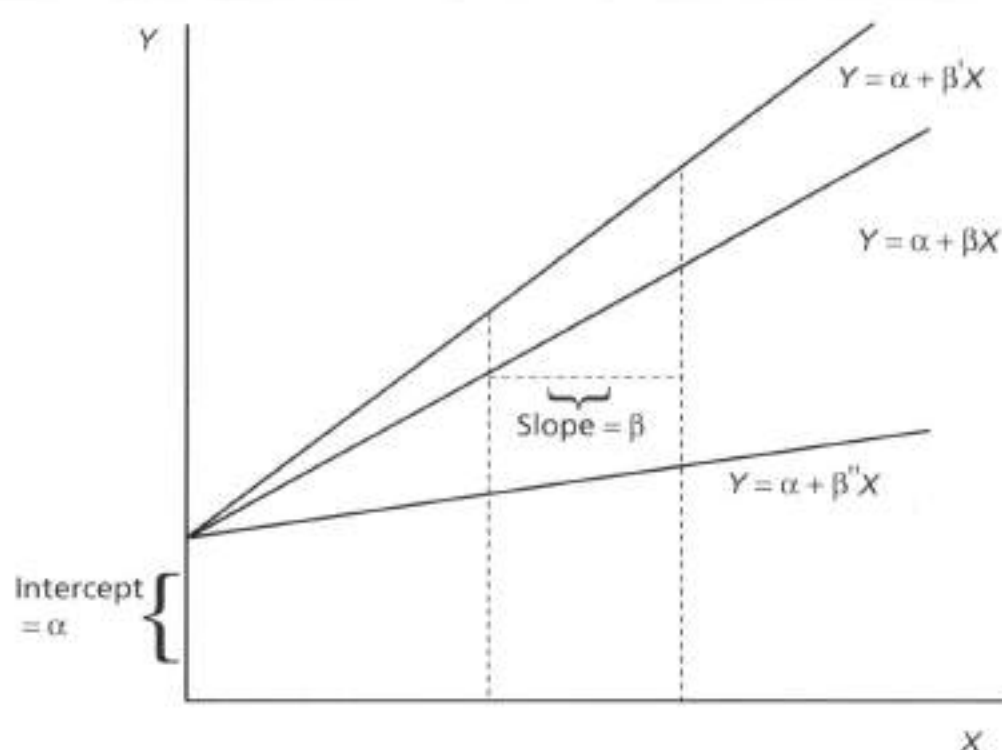
Therefore, we can lay down our criterion for choosing the best line as *the one that gives us the least amount of squared deviation from the observed points*. This is the notion behind the method of least squares.² The equation derived through the method of least squares is called the **regression equation**.

LINEAR TIME TREND: SIMPLE REGRESSION MODEL

The regression model for a straight line is given by the following equation:

$$Y = \alpha + \beta X + \varepsilon \quad (12.1)$$

where Y is the **dependent variable** (the one to be explained), α is the **intercept**, β is the **coefficient for trend** (or slope), X is the **independent variable** (the one that explains the

FIGURE 12.5 Explanation of a Straight-line Equation

dependent variable), and ϵ is the **error of estimation** (or the residual difference between actual Y and estimated $\hat{Y}$).

Figure 12.5 is an explanation of the meaning of a straight-line equation. The intercept term (α) measures the point at which the trend line meets the vertical axis. Quite often, the intercept term has interesting interpretations. For example, in fitting a Keynesian aggregate consumption function (the relation between consumption and income), the intercept term measures the level of consumption one must have even when income is zero. This is the subsistence level of consumption, which presumably must be met either through government transfer (welfare payment) or private charity.

The trend coefficient β measures the magnitude that a one-unit change in the independent variable causes in the dependent variable. Thus, a higher value of β' would signify a steeper line, implying that a small change in the independent variable would cause a great deal of change in the dependent variable. The reverse would be true for a smaller trend coefficient (β'').³ A positive trend coefficient implies an upward slope; a negative coefficient a downward slope; and a zero coefficient implies the absence of any relationship between the dependent and independent variable.

An estimated equation for a straight line contains three components: the coefficient (also known as the constant or parameter) for the intercept, the coefficient for the slope, and the error term:

$$\hat{Y}_i = a + b\hat{X}_i \quad (12.1a)$$

where $\hat{Y}_i$ is the estimated value of Y_i , a is the estimated intercept coefficient, and b is the estimated slope coefficient.

It is important to note the difference between the “true” coefficients and their estimated values. The true values are conceptual in nature and, as such, cannot be known. The best we can hope for is to obtain their closest estimated approximation. To help in understanding this important difference, I have denoted the true values with Greek letters and their estimated values with Roman letters.

The **error term** is the difference (the residual) between the *actual* and the *estimated* values of Y : $e_i = Y_i - \hat{Y}_i$. By replacing $\hat{Y}_i$ with $a + bX_i$ (from equation 12.1a), we get

$$e_i = Y_i - a - bX_i \quad (12.2)$$

As I explained earlier, by the method of least squares we estimate the coefficients of a and b such that we minimize this sum of errors⁴:

$$\text{Minimize } \sum_{i=1}^n (e_i)^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2 \quad (12.3)$$

From the preceding expression (12.3), we can devise the following formula for the estimated coefficient of b :

$$b = \frac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2} \quad (12.4)$$

where $\bar{Y}$ and $\bar{X}$ are the means for the variables Y_i and X_i .

The intercept term is estimated by

$$a = \bar{Y} - b\bar{X} \quad (12.5)$$

Let us consider the following example. Suppose that for the purpose of long-term planning, we would like to estimate the future population of a small town. The population figures for the past five years are shown in Table 12.1. By plotting this series (see Figure 12.6), we can detect the existence of a steady linear trend. We then decide to forecast the future population size using a linear equation form, shown in equation 12.1.

To estimate the coefficients a and b , for a straight-line equation, we need to work out Table 12.2.

TABLE 12.1 Population over Time (in thousands)

Population	Year
5	1
8	2
10	3
15	4
17	5

TABLE 12.2 Calculations for a Straight-line Regression Equation

Population (Y_i)	Year				
	(X_i)	$(Y_i - \bar{Y})$	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$	$(Y_i - \bar{Y})(X_i - \bar{X})$
5	1	-6	-2	4	12
8	2	-3	-1	1	3
10	3	-1	0	0	0
15	4	4	1	1	4
17	5	6	2	4	12
Total		0	0	10	31

Note: $\bar{Y} = 11.0$ and $\bar{X} = 3.0$.

Recalling the formula for calculating b ,

$$b = \frac{\sum(Y_i - \bar{Y})(X_i - \bar{X})}{\sum(X_i - \bar{X})^2}$$

we can calculate the value of b by inserting the numbers from Table 12.2. Thus, we estimate

$$b = \frac{31.00}{10.00} = 3.1$$

Similarly, we can calculate the intercept term a as

$$a = \bar{Y} - (b\bar{X}) = 11 - (3.1 \times 3) = 1.7$$

Therefore, we can write the estimated equation as

$$\text{Population } (\hat{Y}) = 1.7 + 3.1 \text{ Year } (X) \quad (12.6)$$

The plot of the actual data and the trend line are shown in Figure 12.6.

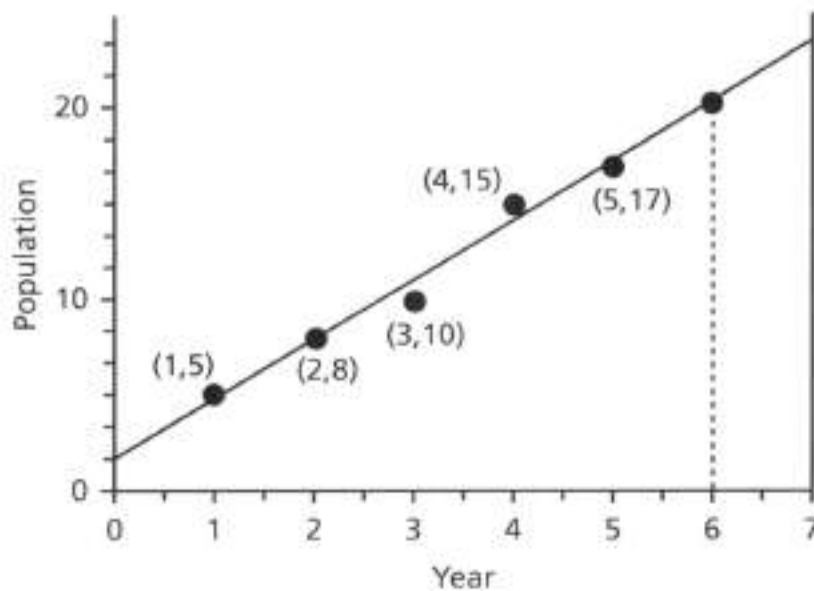
We have plotted the estimated and actual values of population in Figure 12.6. From this derived equation we can project the future population of the city. For example, we can estimate the population for the sixth year (6) as

$$1.7 + (3.1 \times 6) = 20.3$$

Similarly, to obtain the projected figure for the seventh year, we must substitute 7 as the value for the independent variable (X). This arithmetic calculation yields 23.4, the projected population figure for the seventh year.

Accuracy of the Results

Having estimated the trend line, you can legitimately ask, how good are my results? This question has two interrelated parts: First, you may ask whether your model is offering a significant explanation of the variations within the series. Second, you would want to

FIGURE 12.6 Population Trend

explore how confident you feel about your estimation of the coefficients. That is, are the estimated values reflective of the “true” relationship between the dependent and independent variables, or did you obtain a correlation by chance—a spurious correlation?

Recall Figure 12.3. While estimating the relationship between the dependent and independent variables, we can see that we can explain part of the series through estimation, whereas the remainder (the difference between the observed and the estimated values) is unexplained. A perfect explanation explains all the variations within a series. In contrast, in the absence of any relation between the two variables, none of the variations will be explained. Hence, the measure of goodness of fit is the proportion of explained variance within the total variance. Again, since we want to get rid of the problem of having to add positive deviations with negative deviations, we square the deviations and then sum the results.⁵ This ratio is called the **coefficient of determination**, or R^2 , and is written as

$$R^2 = \frac{\text{Squared sum of explained variations}}{\text{Squared sum of total variations}} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \quad (12.7)$$

where Y_i represents the actual values for Y observations, $\bar{Y}$ is the mean of Y , and $\hat{Y}_i$ represents the predicted values of Y_i on the basis of the regression line (calculated using the estimated equation $Y = a + bX_i$).

You may note that when all the Y_i values are correctly predicted, the predicted values are the same as the actual values. In that case, $\hat{Y}_i$ is equal to Y_i , which means that the numerator is equal to the denominator, or the value of R^2 is equal to 1. This is a perfect explanation. When a model cannot explain any variation (or when the variations are completely random), however, the values of $\hat{Y}_i$ form a horizontal line (such as the one shown in the previous chapter in Figure 11.3) that goes through the average value of the distribution. In such a case, $\hat{Y}_i$ is equal to $\bar{Y}$, which makes the numerator equal to 0 and

TABLE 12.3 Calculation of R^2

Population (Y_i)	Year (X_i)	Projected value ($\hat{Y}_i$) ^a	$(\hat{Y}_i - \bar{Y})^2$	$(Y_i - \bar{Y})^2$
5	1	4.8	38.44	36.0
8	2	7.9	9.61	9.0
10	3	11.0	0.00	1.0
15	4	14.1	9.61	16.0
17	5	17.2	38.44	36.0
Total			96.10	98.0

^a The projected values are calculated with the help of the estimated equation 12.6.

consequently R^2 equal to 0. Therefore, the value of R^2 will always fall within the limits 0 (no explanation) and 1 (perfect explanation).

Let us calculate R^2 for our example from Table 12.3. We can calculate R^2 using equation 12.7:

$$R^2 = \frac{96.1}{98.0} = 0.981$$

In other words, our model is predicting slightly more than 98 percent of the variations.

High R^2

There is no universally acceptable answer to the question of what a high R^2 is. Generally speaking, if you are dealing with *economic time-series data*, you are likely to find a high degree of correlation, because most economic variables tend to move in the same direction. In times of prosperity, all the good indicators tend to go up (for example, savings, investment, employment, housing construction). The reverse takes place during economic recession. In an interesting study, Edward Ames and Stanley Reiter showed that even when unrelated time-series data were chosen at random and regressed against one another, R^2 values exceeded 0.5.⁶ The lesson of their study is that since correlation between two variables only measures co-occurrence and does not necessarily establish a causal linkage, a high value of R^2 should not give cause for celebration, unless such a relationship can be backed up by solid theoretical reasoning.

You should also note that for cross-section data, the R^2 values are typically smaller than those for time-series data. Also, because of greater complexity in their interrelationships, correlation coefficients for sociological or political data are often much smaller than those for economic or financial data.

Relevance of the Estimated Coefficients

Now that we have determined how much of an explanation of the data the model is providing, we may want to know the following: how relevant are the *individual coefficients*, and how relevant are the coefficients *taken as a whole*? To answer these questions, we need to look back at the development of theoretical statistics. One of the most remarkable theorems of the entire field of mathematics and mathematical statistics was derived nearly 200 years ago and provides the basis for hypothesis testing. Suppose you

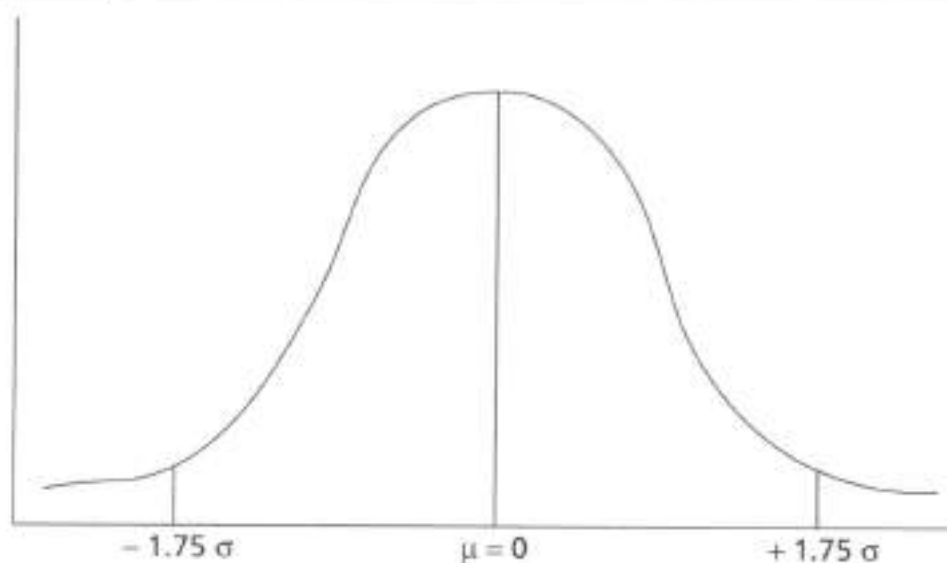
know the average (mean) height of the American male population (μ) and the standard deviation (σ). You are conducting a random survey, whereby you are surveying groups of men and noting their average heights. At every try, you are getting a small sample group together and noting its average height, subtracting it from the **population mean** (μ), and dividing by the **population standard deviation** (σ).⁷ You continue with this process, which symbolically can be written as

$$\frac{(X_i - \mu)}{\sigma} \quad (12.8)$$

As discussed in chapter 7, if you repeat your experiment, the plotted results will approach a normal distribution. This distribution will have a mean equal to 0 and a standard deviation equal to 1. The remarkable aspect of this theorem is that it does not matter what the distribution of the variable X_i is. Unless there are some systematic biases (such as measuring only basketball players to arrive at the average height of the general public), repeating this procedure will give us the normal distribution as the number of samples increases. This theorem is known as the law of large numbers (see chapter 7). By the property of this standardized distribution, we know how much of the values will fall within what range, which makes the derivation of a normal distribution particularly fortuitous. For example, we know that approximately 95 percent of all values will fall within ± 1.96 standard deviations, and nearly 99 percent of all values will fall within ± 2.57 standard deviations. Therefore, in our example, suppose we come across a group of Pygmies with an average height that is more than 1.96 standard deviations below the average for men from the United States. We can then state with a 95 percent level of confidence that the members of this particular sample are not Americans. Let me explain this with the help of Figure 12.7.

You may notice, however, that I have emphasized the term *random*. One of the most important assumptions of a regression model is that the model includes all of the

FIGURE 12.7 Curve of Normal Distribution





A CASE IN POINT

Campaign Contributions and Judicial Impartiality: How Regression Analysis Can Help

Although judges are required to follow the law when deciding cases, there might be enough flexibility in the law to allow for some judicial interpretation. Assuming that a case allows a judge some discretion in decision making, it is generally assumed that judges and justices want to hand down decisions that are consistent with their own personal belief systems. But what if a judge's personal beliefs are inconsistent with the views of a majority of American citizens? Are there any political consequences for rendering unpopular decisions? On the U.S. Supreme Court, to which justices are appointed by the president for life terms, it is unlikely that justices will face any serious ramifications for ruling against the majority. That is not the case, however, on state supreme courts.

Some state high court justices are appointed, but most have set term lengths and keep (or lose) their seats on the bench through popular elections, just like other politicians. And like other politicians, these justices must raise money to fund their campaigns. Most of the money raised by judges comes from special interest groups, many of which have cases before those same judges. If justices' campaign contributors have specific policy objectives and specific rulings they expect justices to deliver, then justices who are reelected may be influenced in their rulings. As such, we can assume that a rational judge has two goals: the judge wants to implement personal policy preferences and wants to get reelected. The question for social scientists is, how do judges balance these two goals?

The question of judicial impartiality came to a head in 2009, when the Supreme Court decided a case that had become a flashpoint in the national debate over potential corruption in state judicial elections caused by big contributors. In a 5-4 decision in the case of *Caperton v. A.T. Massey Coal Company, Inc., et al.*, the Court ruled that "West Virginia Supreme Court Justice Brent Benjamin violated constitutional due process of law when he voted in the dispute after being asked to take himself out because of a conflict of interest."⁸ The perception of judicial impartiality is the cornerstone of our political system. Therefore, the question at hand is of great importance to all of us.

The question is, Is there any empirical proof of campaign money influencing judges' decisions? By creatively using regression analysis, in a series of articles Madhavi McCall provided empirical proof for the question of justice for sale.⁹ In her model, she presented a list of variables that justices consider when making decisions. This is because before we can truly understand the impact

relevant explanatory variables. Therefore, the variation in the observed data that is not explained is caused by random factors, which are distributed normally around the mean of 0. A variable is considered random if its values are determined by the outcome of a "chance experiment." That is, in our example, if we confined our search to a particular ethnic group, or a group of basketball players, then the outcomes are not being determined by chance. Whether the selection is by design or by some unrecognized bias on the part of the researcher, this theorem of the law of large numbers will not apply; the

of campaign contributions on case outcomes, we must control for other factors that might influence judicial choices. For instance, while there may be some flexibility in a justice's ability to interpret the law, the law still matters and is most likely a factor in judicial decisions. Additionally, some judges may choose to follow more closely what campaign contributors want in the months immediately before they face reelection in the hope of generating more contributions, suggesting that the timing of a case during a justice's term also matters. After getting reelected, no longer immediately needing popular votes or campaign contributions, a rational judge may decide that he or she now has an opportunity to decide cases based only on his or her personal view of the law, suggesting that a justice's personal policy preferences are important. Using regression analyses, McCall tested what impact campaign contributions, case timing, the law, and a justice's personal policy preferences have on that justice's eventual vote and determined which factors, when they are all considered together, exert the strongest influence on the justice's rulings.

For instance, when regression was applied to the voting behavior of Texas's elected state supreme court justices, McCall found that the justices' voting decisions appeared to be strongly influenced by campaign contributions and that the Texas justices appeared to be very strategic in their decisions. Immediately before elections, the justices voted with the popular majority so that they could garner sufficient votes to win the elections. However, during the campaign season (several months of fund-raising and campaigning before the actual elections), the justices were more inclined to vote with campaign contributors. This was true even when a campaign contributor held a fundamentally different legal position than the justice typically supported, suggesting that during the campaign season, justices put aside their own personal policy positions and instead favored the positions of campaign contributors. And of course, certain aspects of the law restricted judicial choices. Ultimately, however, the model indicates that during the campaign season, the most significant influence on judicial behavior is the presence or absence of a campaign contribution. The use of regression in this type of study allows for a better understanding of the determinants of judicial behavior.

Discussion Points

1. On the basis of the findings of these studies, what kind of policy would you propose to ensure judicial impartiality? For instance, should judges be appointed for life or, as in some countries, rise through the ranks of civil service?
2. Can you think of and propose testable hypotheses in other areas of policy inquiry?

resulting distribution will not approximate a normal distribution, regardless of the number of experiments. As a result, none of the results described below will hold true.

The Significance of Individual Coefficients

As can be seen in Figure 12.7, a group average sufficiently different from the mean may indicate that the two populations are indeed different. Hence, this theorem serves as the basis for testing the validity of hypotheses. For our purposes, we can use this result to test

the null hypothesis that our coefficients are results of chance and, in fact, there is no relationship between the dependent and the independent variables. For example, the level of statistical significance for our estimated trend coefficient (b) can be calculated as

$$\frac{b - \beta}{\sigma} \quad (12.9)$$

where b is the estimated coefficient, β is the true coefficient, and σ is the true standard deviation of the distribution. Since we are testing the null hypothesis that there is no correlation between the dependent and independent variables, we set β equal to 0. Also, since we do not know what the true standard deviation is for population, we approximate it within the estimated standard error of coefficient, which we may call SE_b . Therefore, equation 12.9 can be written as

$$t_b = \frac{b}{SE_b} \quad (12.10)$$

Similarly, to test the statistical significance of the coefficient a , we need to calculate the t ratio given by

$$t_a = \frac{a}{SE_a} \quad (12.11)$$

From equations 12.10 and 12.11, we can see that the larger a coefficient is in relation to its standard error of estimate, the farther out the calculated ratio will be on our normal distribution chart. The farther out it is, the more convinced we will be that the coefficients are not equal to 0. Therefore, the value of this ratio will tell us that the coefficients are **statistically significant**.

Before we get into the process of actually testing the null hypothesis regarding the validity of the coefficients, I would like to repeat a few important properties of hypothesis testing from our discussion in chapter 7. First, as the name suggests, by the law of large numbers, the normal distribution is achieved only when the number of observations becomes extremely large. In such cases one can use the Z table to test hypotheses. However, when the number of observations is relatively small, an approximation of the normal distribution is used to ascertain the critical values. This is called **Student's t distribution**, named after a statistician who wrote under the pseudonym Student.

Since we do not know the true standard deviation of an estimator (an estimated coefficient), we need to approximate it with our sample results. Let us take a concrete example. In our population projection, we estimated two coefficients: $a = 11.0$ and $b = 3.1$. On the basis of that result, we can use the following sample estimate of the true variance σ^2 :

$$SE^2 = \frac{\sum e_i^2}{n - 2} = \frac{\sum (Y_i - a - bX_i)^2}{n - 2} \quad (12.12)$$

where SE is the standard error of estimate.

You can see that the numerator of equation 12.12 is the sum of squared errors of estimate. This sum is being divided by $n - 2$ (the number of observations minus 2) to approximate the true variance.¹⁰

With an estimate of $\sum e_i^2$, we can estimate the variance associated with the estimated coefficients, a and b . The respective standard errors are calculated as

$$SE_a = \sqrt{SE^2 \frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2}} \quad (12.13)$$

and

$$SE_b = \sqrt{\frac{SE^2}{\sum (X_i - \bar{X})^2}} \quad (12.14)$$

From the preceding formulas, we can see that to calculate the standard errors of the coefficients, we need to add a column to Table 12.3 (see Table 12.4).

From equation 12.12, the estimate of $SE^2 = 1.9/3 = 0.63$. Substituting the numbers in equations 12.13 and 12.14, we derive the estimated standard error of coefficients a and b as

$$SE_a = \sqrt{0.63 \left(\frac{55}{5 \times 10} \right)} = 0.83$$

and

$$SE_b = \sqrt{\frac{0.63}{10}} = 0.25$$

Substituting the respective values in equations 12.10 and 12.11, we get

$$t_a = \frac{1.7}{0.83} = 2.05 \quad (12.13a)$$

and

$$t_b = \frac{3.1}{0.25} = 12.4 \quad (12.14a)$$

For an interpretation of these results, we can resort to Figure 12.7. We noted that if the estimated coefficient is located farther away from its critical value, we can be reasonably certain that the estimated value is different from the true value proposed by our null hypothesis.

These values (t_a and t_b) are distributed as t , with $n - 2$ degrees of freedom (see definition in chapter 7). Since we have five observations, our degrees of freedom are $5 - 2 = 3$. By checking a t table for 3 degrees of freedom, we see that for a .05 level of significance (a 95 percent confidence level), the t ratio is equal to 2.353. Therefore, if the ratios are greater than this number, we can state with a great deal of confidence that our estimated coefficients are indeed significantly different from zero. In other words, there is a significant relationship between the dependent and the independent variables.

Because the t value of the intercept term (2.1) is less than the required critical value (2.353), we cannot reject the null hypothesis that the intercept is equal to 0. In contrast, because the t ratio for the slope (b) is greater than this number, we can be sure that this number is statistically significant.

TABLE 12.4 Calculations for Tests of Significance

Population (Y_i)	Year (X_i)	Projected value ($\hat{Y}_i$)	$(Y_i - \hat{Y}_i)^2$	$(X_i - \bar{X})^2$	X_i^2
5	1	4.8	0.04	4	1
8	2	7.9	0.01	1	4
10	3	11.0	1.00	0	9
15	4	14.1	0.81	1	16
17	5	17.2	0.04	4	25
Total			1.90	10	55

Presentation of Estimation Results

Once a regression equation is estimated, the results must be presented to the policy-makers. There are of course many ways of presenting this information. I suggest the following:

$$\text{Population} = 1.7 + 3.1\text{Year}$$

$$t = (2.037)(12.32)^*$$

$$R^2 = 0.98$$

$$\text{Adjusted } R^2 = 0.97$$

$$n = 5$$

$$F_{(1,3)} = 151.74^*$$

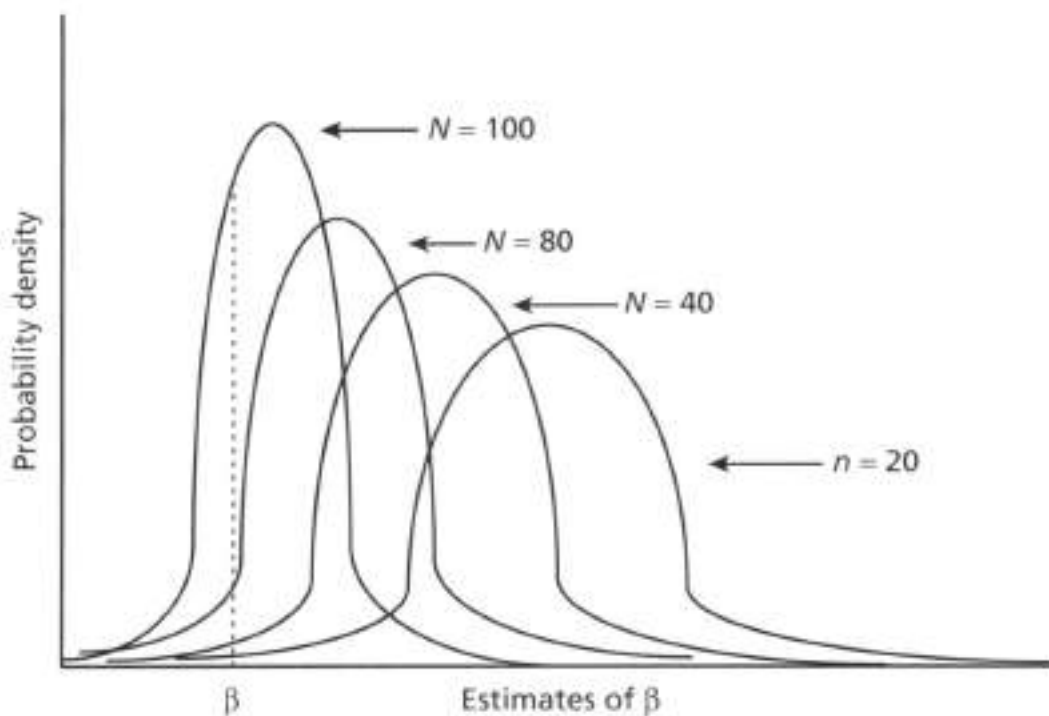
(12.15)

*Significant at the .05 level

The numbers within the parentheses in the estimated equation 12.15 are the respective t values for the individual coefficients. The asterisks point out the significant t values and F value. You may notice that the t value for the slope coefficient and the F statistic are significant, but the t value for the intercept term is not significant at the .05 level. An estimated equation with a high R^2 and significant slope coefficients but an insignificant intercept may seem a bit confusing to you. It may happen for two reasons. First, it may be that the “true” intercept is too close to 0, and thus the regression equation is unable to determine its estimated value. Or it may mean that you have left out some important independent variable in your construction (technically, “specification”) of the model.

The Number of Observations

A series with five observations occurs in my example of population growth. You may ask, what is the minimum number of observations I need to establish a proper relationship between the dependent and the independent variables? Remembering the law of large numbers, we can see that as the number of observations increases, so does the accuracy of the estimated coefficients. In Figure 12.8 I have shown how the estimated

FIGURE 12.8 Increasing Accuracy of Estimated Coefficients

coefficients tend to converge on the “true” coefficients as the number of observations becomes very large (and, consequently, the sample distribution approaches a normal distribution). You may also note that with a larger number of observations, not only do the estimated coefficients come closer to the actual number, but their deviation gets smaller. That is, we are increasingly confident that our estimated **regression coefficients** represent their “true” values. Finally, no number can be considered sufficiently large to estimate the “true” value of the coefficient, so I have drawn the estimated coefficient b , with 100 observations, close to its true value of β , but they are still not equal.

In sum, because there is no definition of a “very large number,” as a rule of thumb you should not have fewer than ten observations in a regression model. My examples of the estimated equation have fewer than ten observations solely for the ease of computation.

TREND CHANGES: BUILDING MULTIPLE REGRESSION MODELS

In a straight-line model we hypothesize that the trend has remained unchanged during the period under study. However, it is entirely possible that the data may reflect a changing pattern. This change can be abrupt or can be part of a gradual process. An abrupt change in trend can result from sudden catastrophic events such as war, the changing of a law, a significant invention resulting in a change in technology,

or a change in policy resulting from a change in political leadership. These events introduce a **qualitative change** in a series. For example, if we note the trend of per capita GDP in a country such as Iraq, we may notice dramatic downward shifts in trend following Iraq's two devastating wars with Iran and the coalition led by the United States. Similarly, several data series (such as national debt) for the United States will show signs of abrupt changes brought about by significant wars. A study of immigration patterns will clearly show the results of changes in the law. Similarly, the invention of many modern drugs, such as penicillin, caused shifts in the trends of infant mortality in the late 1940s and 1950s.

Abrupt Changes in Trend

Suppose we are studying the trend in per capita expenditure by the local government in a community that incorporated itself as a new city in 1996. The achievement of its city status has enabled it to access more state and federal grants, which has caused a vertical shift in the series. Table 12.5 provides the series of per capita government expenditures.

This abrupt change in trend is shown in Figure 12.9. You may notice that the vertical shift that takes place during the middle of the series (in 1996) does not affect the trend pattern itself.

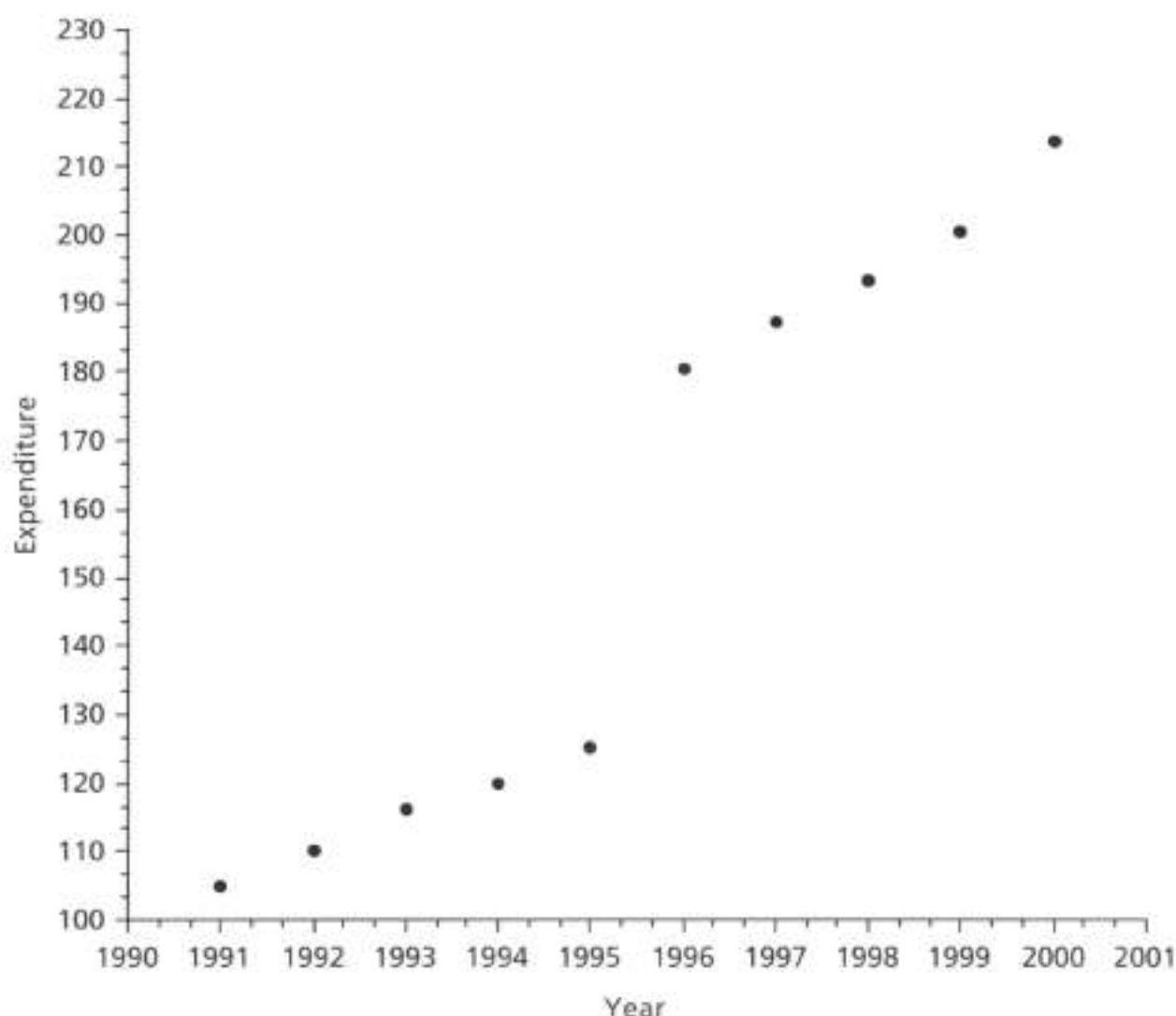
These sorts of shifts are typical of series in which an extremely important external event has taken place or there is a large gap in the data set. It is clear from this figure that a broken trend line would produce fewer errors of estimation than would an unbroken straight line fitted through the entire series.

For a better visual presentation, I have replotted the data shown in Figure 12.9 (see Figure 12.10). In this figure, we can see the result of a vertical shift. We can account for this vertical shift if we draw a trend line based on our pre-incorporation (before 1996) data and then add a constant to the estimated values for the post-incorporation data (1996 and after). This can be done if we have *two* intercept terms, one for before and one for after the incorporation. We can accomplish this with a trick: we can introduce what is known as an **intercept dummy variable**.

TABLE 12.5

Per Capita Federal Grant Expenditures

Year	Expenditure (constant \$)
1991	105.0
1992	110.0
1993	116.0
1994	120.0
1995	125.0
1996	180.0
1997	187.0
1998	193.0
1999	200.0
2000	213.0

FIGURE 12.9 Abrupt Shift in a Series

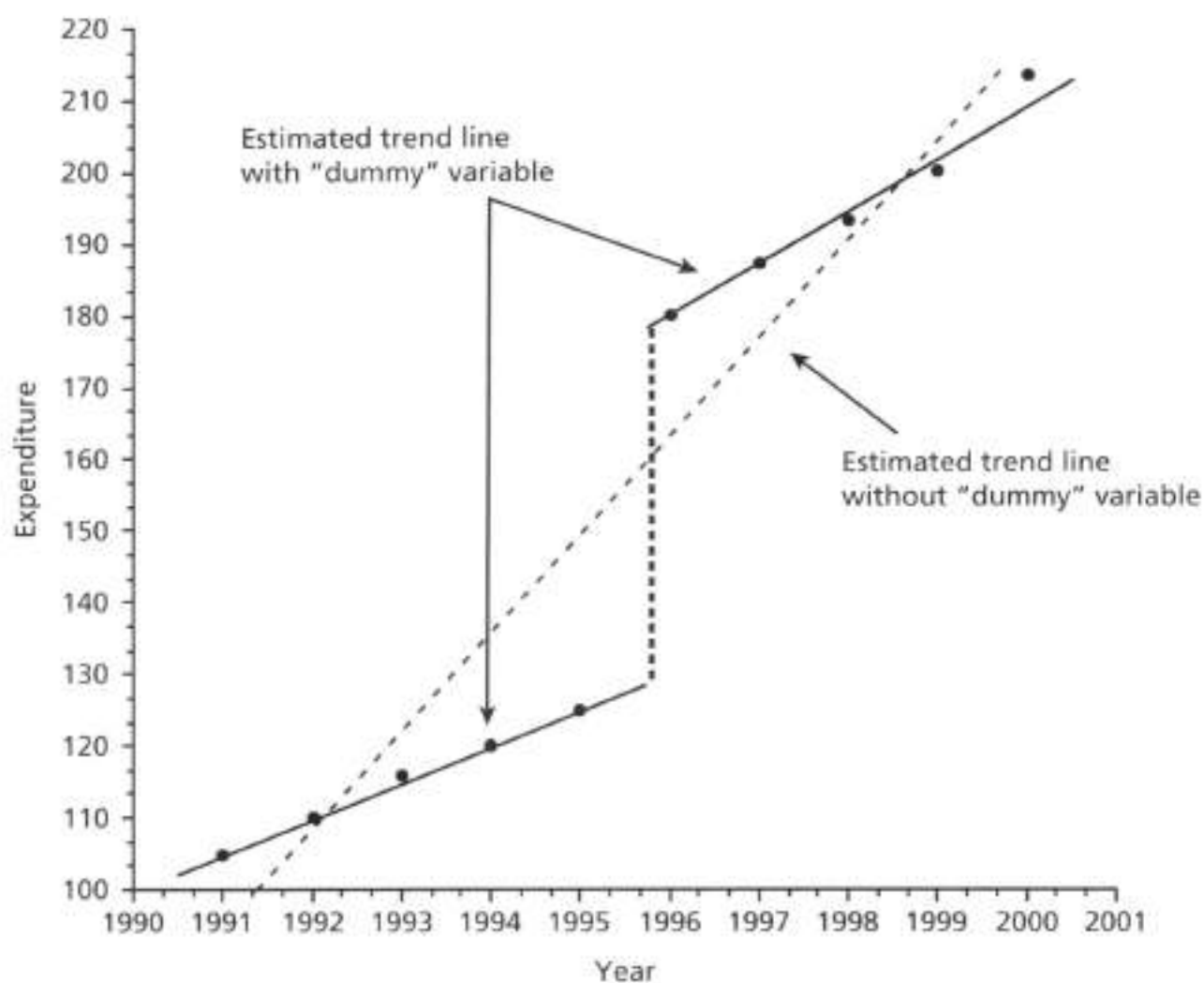
An intercept dummy introduces an additional intercept term to account for this kind of a shift. This takes the form of adding an extra independent variable, which has a value of 0 for the observations prior to the time of shift and 1 for periods afterward. Therefore, the equation to be estimated is written as

$$Y_i = a + a' + bX_i + e_i \quad (12.16)$$

where a' is the dummy variable, which is 0 for years prior to the shift (1991–1996 in our example) and 1 for the years afterward (1996–2000).

Let us proceed with the example of our newly incorporated city. Since we want to introduce an additional independent variable, a' , we need to add a new column to Table 12.5, as seen in Table 12.6.

You may realize now that as we introduce an additional independent variable, a' , we will be using multiple regression instead of simple regression. The estimation of the coefficients of multiple regression is a bit too complicated to be calculated by hand. Even

FIGURE 12.10 Shift in Expenditure Pattern**TABLE 12.6** Setting Up an Intercept Dummy Variable

Year	Expenditure (constant \$)	Dummy (a')
1991	105.0	0
1992	110.0	0
1993	116.0	0
1994	120.0	0
1995	125.0	0
1996	180.0	1
1997	187.0	1
1998	193.0	1
1999	200.0	1
2000	213.0	1

though generations of students of statistics and econometrics were subjected to this tortuous exercise, the proliferation of computers has made the estimation rather simple. Therefore, you will want to use suitable statistical software to derive the coefficients for this equation.

The use of this dummy variable gives us an excellent predictive model, as shown by the results of the estimated equation¹¹:

$$\begin{aligned} \text{Expenditure} &= -12,739.65 + 6.45 \text{ Year} + 47.15 \text{ Dummy} \\ t &= (-9.768)^* \quad (9.857)^* \quad (12.543)^* \\ R^2 &= 0.996 \\ \text{Adjusted } R^2 &= 0.995 \\ n &= 10 \\ F_{(2,7)} &= 968.7^* \end{aligned} \tag{12.17}$$

*Significant at the .05 level

where the dummy variable equals 0 for years prior to 1996 and 1 for 1996 and afterward.

I have placed the calculated t ratios under the corresponding regression coefficients. As you can tell from this equation, we are explaining 99.5 percent of the total variance in the series. The dummy variable tells us that as a result of incorporation, an average resident of the city gained \$47.15 in government funding. The coefficient for the variable Year tells us that every year, the federal contribution has gone up by \$6.45 per city resident. If we want to predict the level of per capita government expenditure for 2001, we will have to write the equation as follows:

$$\begin{aligned} \text{Projected expenditure (2001)} &= -12,739.65 + 6.45 \times 2001 \\ &\quad + 47.15 \times 1 = 213.95 \end{aligned}$$

To estimate values for years prior to 1996, we will have to set the dummy variable equal to 0, and therefore the equation will contain only the intercept and slope terms. Thus, we may estimate the value for 1993 as follows:

$$\begin{aligned} \text{Estimated expenditure (1993)} &= -12,739.65 + 6.45 \times 1993 \\ &\quad + 47.15 \times 0 = 115.2 \end{aligned}$$

I have plotted the predicted versus the actual values of expenditure in Figure 12.10. The estimated line with the dummy variable has been drawn as a solid line, whereas the straight-line equation has been drawn as a dotted line. You may notice the close approximation of the actual values as a result of the use of the dummy variable. For example, if you estimate the equation without the dummy variable, you get the following result:

$$\begin{aligned} \text{Expenditure} &= -26,971.81 + 13.59 \text{ Year} \\ t &= (-9.26)^* \quad (9.31)^* \\ R^2 &= 0.915 \\ \text{Adjusted } R^2 &= 0.905 \\ n &= 10 \\ F_{(2,7)} &= 86.67^* \end{aligned} \tag{12.17a}$$

*Significant at the .05 level

Although this equation looks excellent on its own, the use of the dummy significantly improves the quality of the estimated equation.¹²

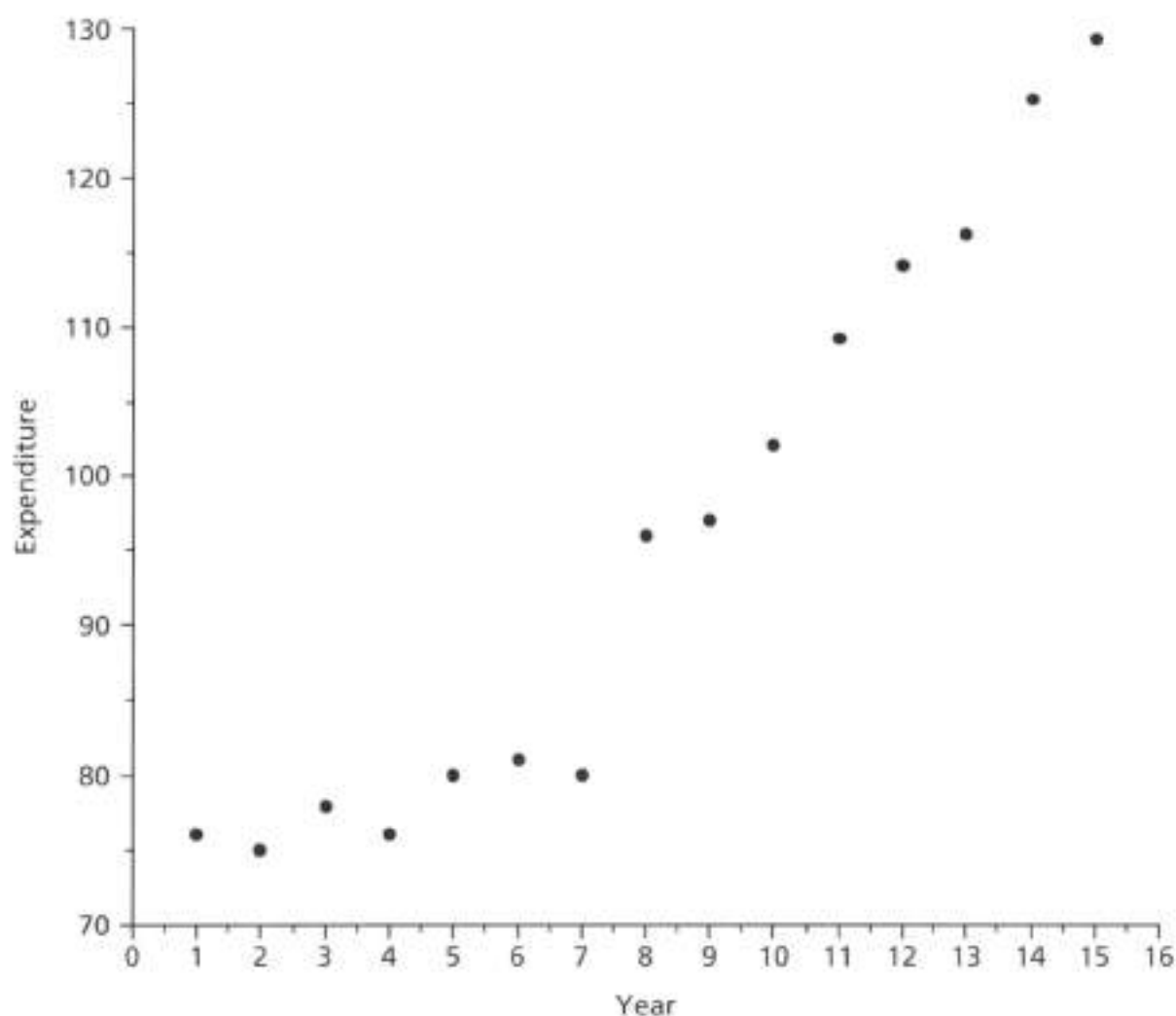
Abrupt Changes in Slope

In the previous example, we discussed the situation in which there is a parallel shift in the trend line. In reality there can also be an abrupt change in the slope. This change is more dramatic because this new situation has altered the entire direction of the trend line. There are many examples of this abrupt change in trend. The discovery of gold in California in 1849 caused a huge population influx; the initiation of the Great Society program in the 1960s caused a sudden increase in the trend of social expenditures; the successful launching of the Soviet Union's *Sputnik* satellite caused a big increase in expenditures for NASA for a decade; during the presidency of Ronald Reagan, federal assistance to state and local governments took an abrupt change in the opposite direction. Let us consider a hypothetical example of government expenditures on drug rehabilitation programs for a period of fifteen years. Assume that because of heightened awareness and a change in political leadership during the eighth year of the study, society decided to spend increasing amounts on drug rehabilitation programs (see Table 12.7 and Figure 12.11).

The plot in Figure 12.11 makes it clear that the trend pattern changed after the seventh year. Whereas for the first seven years of the study, government expenditure was increasing at a slow rate, the rate of change accelerated greatly from the eighth year on. This change in trend cannot be accommodated simply by adding an intercept dummy. Therefore, in our projection of trend, we need to account for this change. This abrupt change can be captured adequately by using dummy variables for the two different slopes. Imagine holding a stick at an angle approximating the trend shown at the beginning of a data series. We can visualize the change in trend if we break the stick at the point at which the change took place. To split the series into two distinct trend lines, we introduce two dummy variables, D_1 and D_2 , with D_1 having the value 1 for all the points

TABLE 12.7 Expenditures on Drug Rehabilitation Programs

Year	Expenditure ($\times \$1,000$)
1	76
2	75
3	78
4	76
5	80
6	81
7	80
8	96
9	97
10	102
11	109
12	114
13	116
14	125
15	129

FIGURE 12.11**Plot of Trend in Government Expenditures for Drug Rehabilitation**

for which the old trend applies and 0 for the rest. D_2 has the values 0 and 1 in the reverse order. Then, by multiplying the independent variable by these two dummy variables, we can create two new sets of independent variables. By running a multiple regression on the dependent variables against these two newly created variables, we can estimate the equation containing two different slopes (Table 12.8).

Using the data shown in Table 12.8, we estimate the trend equation for government expenditure, which yields the following results:

$$\begin{aligned}
 \text{Expenditure} &= 70.61 + 1.663(D_1 \times \text{Year}) + 3.568(D_2 \times \text{Year}) \\
 t &= (27.5)^* \quad (2.8)^* \quad (15.0)^* \\
 R^2 &= 0.974 \\
 \text{Adjusted } R^2 &= 0.969 \\
 n &= 15 \\
 F_{(2,12)} &= 222.75^*
 \end{aligned}
 \tag{12.18}$$

*Significant at the .05 level

TABLE 12.8 **Setting Up Slope Dummy Variables**

Year	Expenditure ($\times \$1,000$)	D_1	D_2	$D_1 \times \text{Year}$	$D_2 \times \text{Year}$
1	76	1	0	1	0
2	75	1	0	2	0
3	78	1	0	3	0
4	76	1	0	4	0
5	80	1	0	5	0
6	81	1	0	6	0
7	80	1	0	7	0
8	96	0	1	0	8
9	97	0	1	0	9
10	102	0	1	0	10
11	109	0	1	0	11
12	114	0	1	0	12
13	116	0	1	0	13
14	125	0	1	0	14
15	129	0	1	0	15

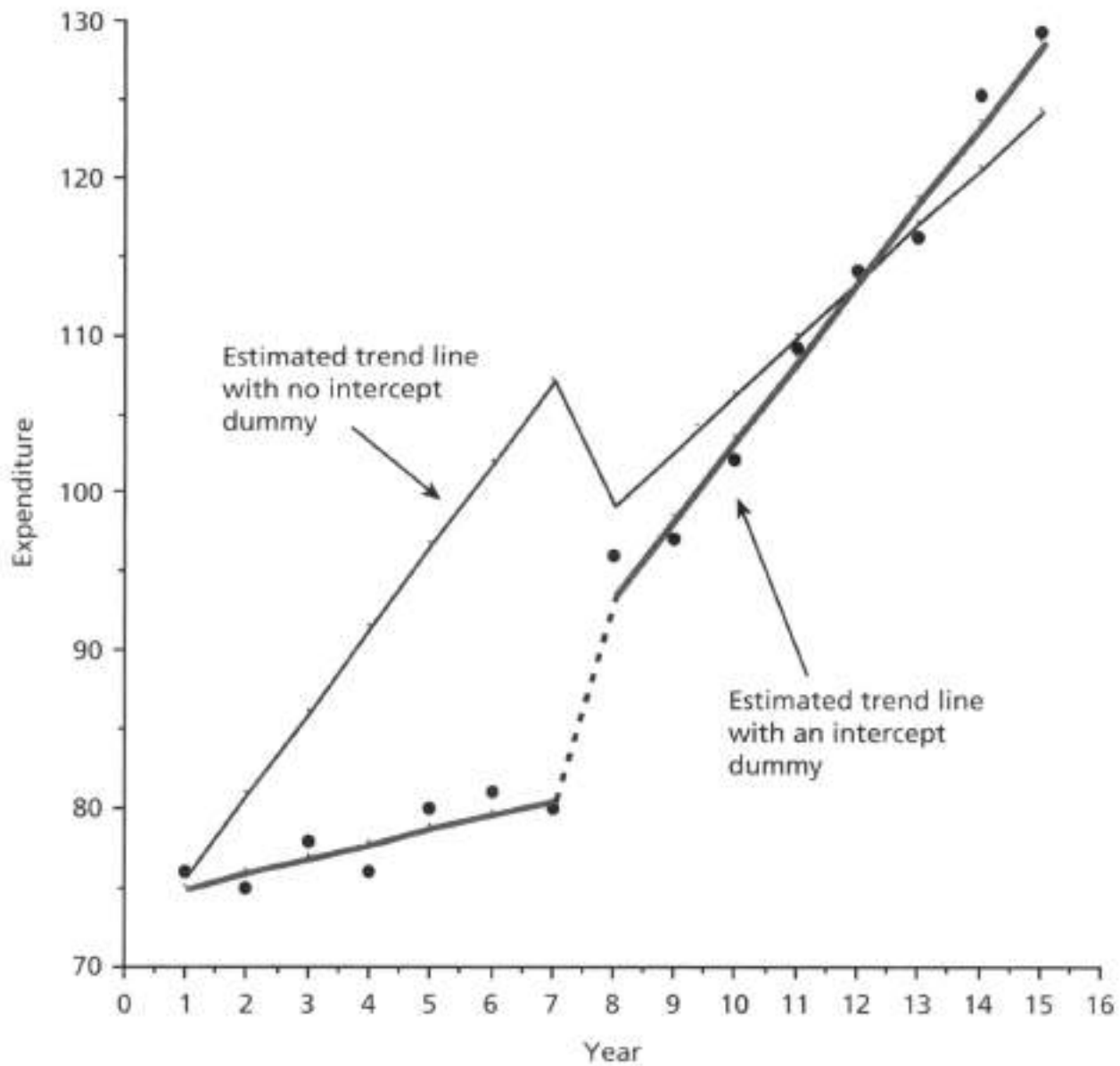
From this equation, we can see that the trend coefficient for the second period (3.568) is 2.15 times larger than that for the first period (1.663). This equation can be used to estimate the amount of government expenditure for every year within the study and can be used to predict the values for future time periods.

If you estimate an equation with a **slope dummy**, you may run into a problem when you attempt to forecast. If you estimate an equation with two slope dummies to account for the two phases of a trend, you may have to include an intercept dummy in your model. If you do not include an additional intercept dummy, the model will force each slope to compromise on a common intercept, which will make both the slope lines deviate from their true positions. If, however, you allow for the slope dummies to have their own intercepts, the estimated results will be far superior.

Consider once again the estimated equation (12.18). We can significantly improve the model by including an intercept dummy, D_1 . The model is estimated as

$$\begin{aligned}
 \text{Expenditure} &= 53.77 + 20.51D_1 + 0.929(D_1 \times \text{Year}) + 4.976(D_2 \times \text{Year}) \\
 t &= (18.05)^* (6.24)^* \quad (2.98)^* \quad (19.58)^* \\
 R^2 &= 0.994 \\
 \text{Adjusted } R^2 &= 0.993 \\
 n &= 15 \\
 F_{(2,12)} &= 630.47^*
 \end{aligned}
 \tag{12.19}$$

*Significant at the .05 level

FIGURE 12.12 Comparison of Estimated Trend Lines

If you compare the estimated results of equation 12.18 with those of equation 12.19, you will notice that all the statistics have improved. However, the real improvement can be seen when the predicted values based on the two equations are plotted side by side against the actual values. In Figure 12.12, I have drawn the predicted values of actual expenditure on the basis of the two models. You can see that the line with a dummy variable is clearly superior (closer) to the actual values. You may also note that the slope of the second phase (4.976) is nearly 5.4 times the slope of the first phase (0.929). I have presented the estimated data in Table 12.9.

TABLE 12.9**Comparison of Estimated Data**

Actual observation	Estimated value (without intercept dummy, equation 12.18)	Estimated value (with intercept dummy, equation 12.19)
76	72.27	75.21
75	73.94	76.14
78	75.60	77.07
76	77.26	78.00
80	78.93	78.93
81	80.59	79.86
80	82.25	80.79
96	99.16	93.58
97	102.72	98.56
102	106.29	103.54
109	109.86	108.51
114	113.43	113.49
116	116.99	118.46
125	120.56	123.44
129	124.13	128.42

GRADUAL CHANGES IN TREND: ESTIMATION OF NONLINEAR TRENDS

Many data series relating to society and the economy can be approximated by a straight line. When a trend is fundamentally nonlinear, however, the use of a straight line will cause severe problems in forecasting. Many nonlinear forms can be captured by regression equations (thanks to advances in computer technology) without a great deal of computational difficulty. Some nonlinear trends can be estimated easily by transforming the data, in which case we do not need to use any special nonlinear estimation techniques (which are beyond the scope of this book). Let us consider a few examples.

Polynomial Forms

A polynomial form expresses a dependent variable as a function of a number of independent variables. Some of these independent variables may be raised to powers greater than 1. The degree of a polynomial is known by the highest power among the independent variables. Thus, a quadratic form, expressed as $Y_i = a + b_1X_i + b_2X_i^2$, is called a second-degree polynomial, the equation $Y_i = a + b_1X_i + b_2X_i^2 + b_3X_i^3$ is called a third-degree polynomial, and so on.

It is common to find examples of a quadratic relationship in nature. For example, as we apply fertilizer to plants, they grow at a rapid rate, and we get more flowers, fruits, and vegetables. However, after a certain point, the application of more fertilizer damages their growth. Or, after an initial period of decrease in the cost per flower, we increase our scale of operation. But at a certain point, inefficiency creeps in and unit costs start

TABLE 12.10**Budget Deficit as a Percentage of GDP (five-year average)**

Years	Percentage of GDP	Time
1970–1974	1.5	1
1975–1979	2.8	2
1980–1984	3.9	3
1985–1989	5.4	4
1990–1994	4.9	5
1995–1999	1.6	6

climbing. As another example, if we start consuming something we highly desire, our satisfaction goes up, but after a point of saturation, we tend to lose interest. These are all examples of quadratic relationships, as discussed in chapter 11. A quadratic form can be U shaped or inverted U shaped.

The quadratic equation is specified as

$$Y_i = a + b_1X_i - b_2X_i^2 \quad (12.20)$$

for a U-shaped relationship and

$$Y_i = a - b_1X_i + b_2X_i^2 \quad (12.21)$$

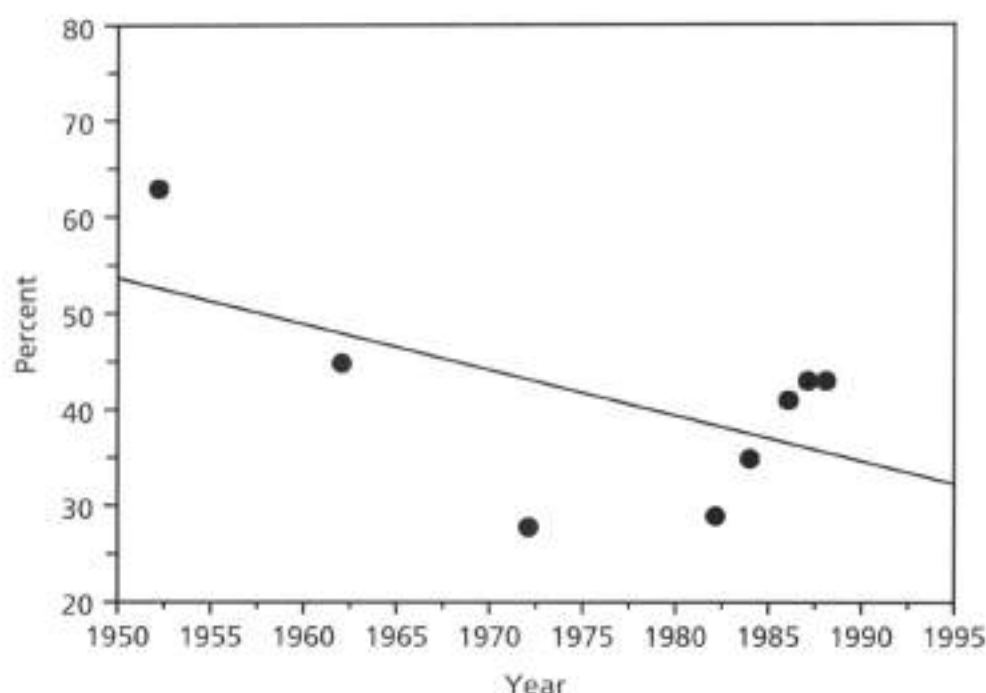
for an inverse U-shaped relationship.¹³

Let us see how these nonlinear forms are estimated. Consider the data series for the federal budget deficit as a percentage of GDP shown in Table 12.10. Note that I have used five-year averages. You may also note that I have replaced the independent variable with a new column, “Time.”

During the first twenty years, between 1970 and 1985, the federal budget deficit as a percentage of U.S. GDP showed a trend of steady increase. However, this trend reversed itself in the next decade. Therefore, if we try to fit a straight-line trend, the approximation of the actual data will be rather poor. This is demonstrated when a linear estimation gives us the following result:

$$\begin{aligned} \text{Percentage deficit} &= 2.52 + 0.237 \text{ Time} \\ t &= (1.52)(0.56) \\ R^2 &= 0.072 \end{aligned} \quad (12.22)$$

Since the critical value for t at a .05 level of significance with 5 degrees of freedom is 2.05, we cannot reject the null hypothesis that the coefficients are in fact equal to 0. As you can clearly see in Figure 12.13, the line fitted through the data does not give us a good prediction. The situation changes dramatically, however, when we try a quadratic

FIGURE 12.13**Publicly Held Federal Debt as a Percentage of GDP:
Linear Estimation**

Source: *Economic Report of the President* (Washington, D.C.: U.S. Government Printing Office, 1989).

fit. To do that, we need to add one more column of independent variables, "Time²" (see Table 12.11).

Now that we have made the necessary data transformation, we can estimate the desired equation. In the same way we estimated the straight-line equation, we can use our computer software to estimate the three coefficients (a , b , and c). In this case, we will have to indicate that we have two independent variables (Year and Year²) instead of one.

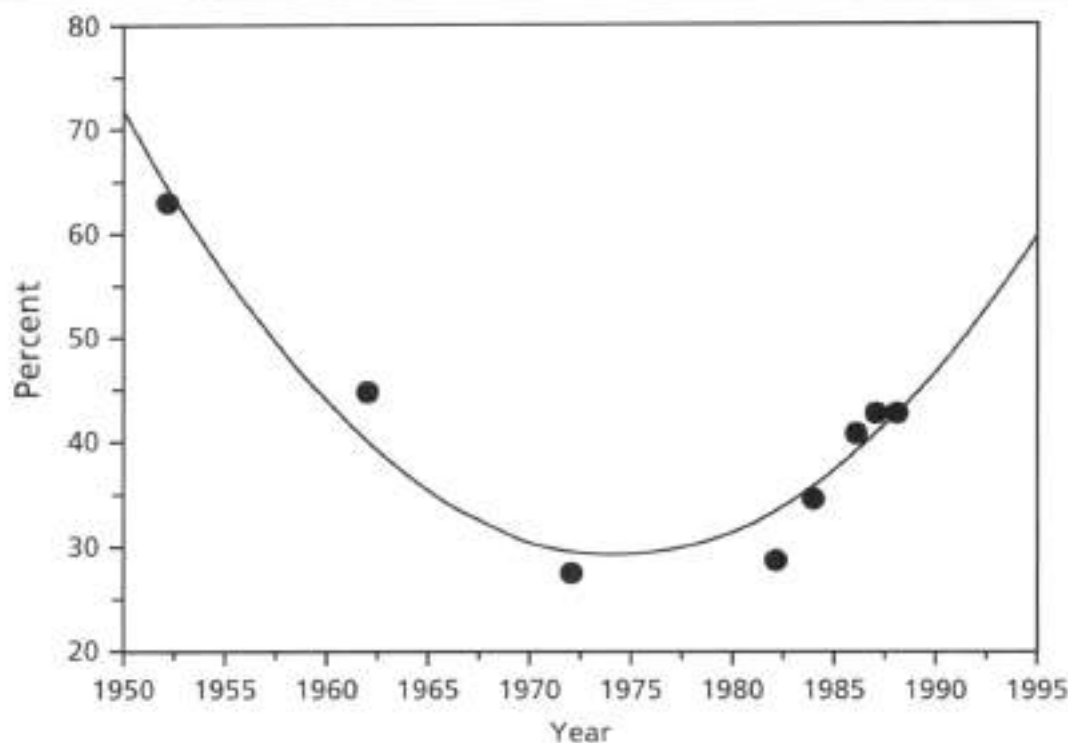
The estimated values of the coefficients are as follows:

$$\begin{aligned} \text{Percentage deficit} &= -2.38 + 3.91 \text{ Time} - 0.525 \text{ Time}^2 \\ t &= (-1.481)(3.72)^* \quad (-3.57)^* \\ R^2 &= 0.823 \end{aligned} \quad (12.23)$$

*Significant at the .05 level. The critical value for t at a .05 level of significance is 2.0.

TABLE 12.11**Transformation of Data for the Estimation of a
Quadratic Equation Form**

Year	Percentage of GDP	Time	Time ²
1970–1974	1.5	1	1
1975–1979	2.8	2	4
1980–1984	3.9	3	9
1985–1989	5.4	4	16
1990–1994	4.9	5	25
1995–1999	1.6	6	36

FIGURE 12.14**Publicly Held Federal Debt as a Percentage of GDP:
Quadratic Estimation**

As you can see from Figure 12.14, the estimated quadratic equation (12.23) is clearly superior to the straight-line equation (12.22). From this equation, we can estimate the value for 1982 to be 29.81 percent, much closer to the actual figure.

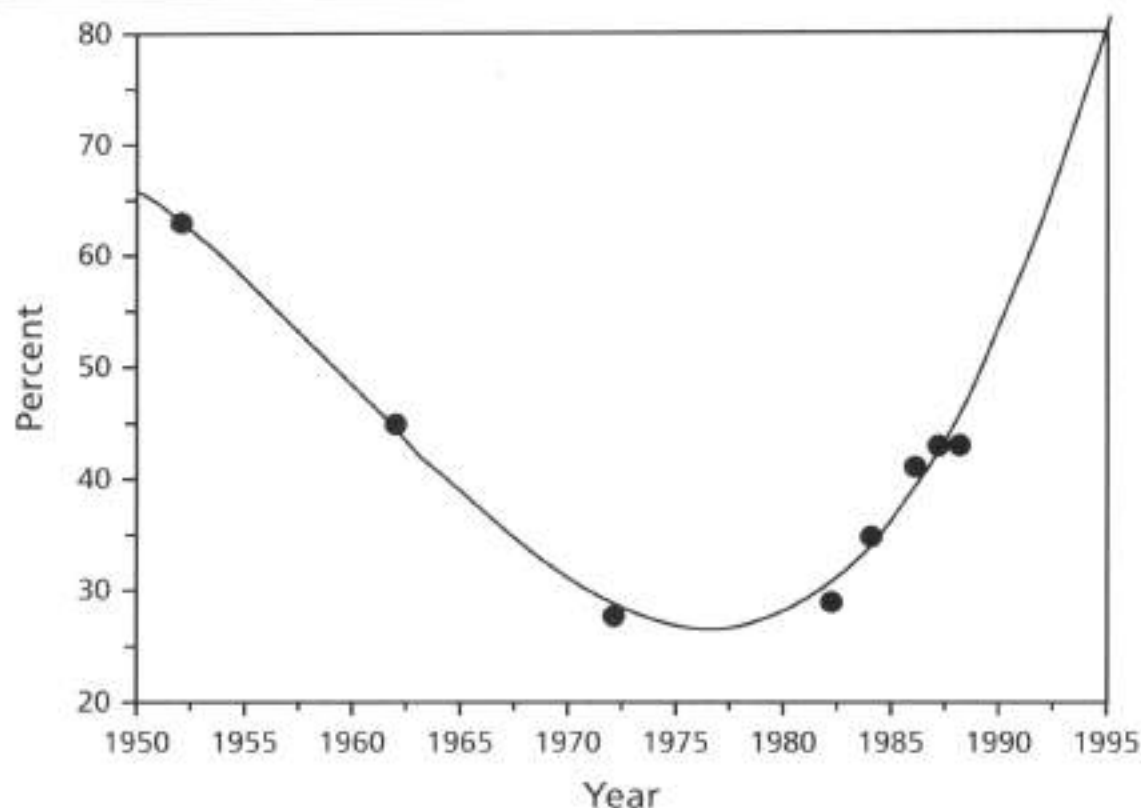
Higher Order Polynomials

Although the regression results indicate that the quadratic form (equation 12.20) seems to be a definite improvement over the linear form (equation 12.17), can we not improve the results even more by using a third-degree polynomial? If we run a third-degree equation, we get the following result:

$$Y = 17,704,000 + 271,050 \text{ Year} - 13.83 \text{ Year}^2 + 3.0 \text{ Year}^3 \quad (12.24)$$

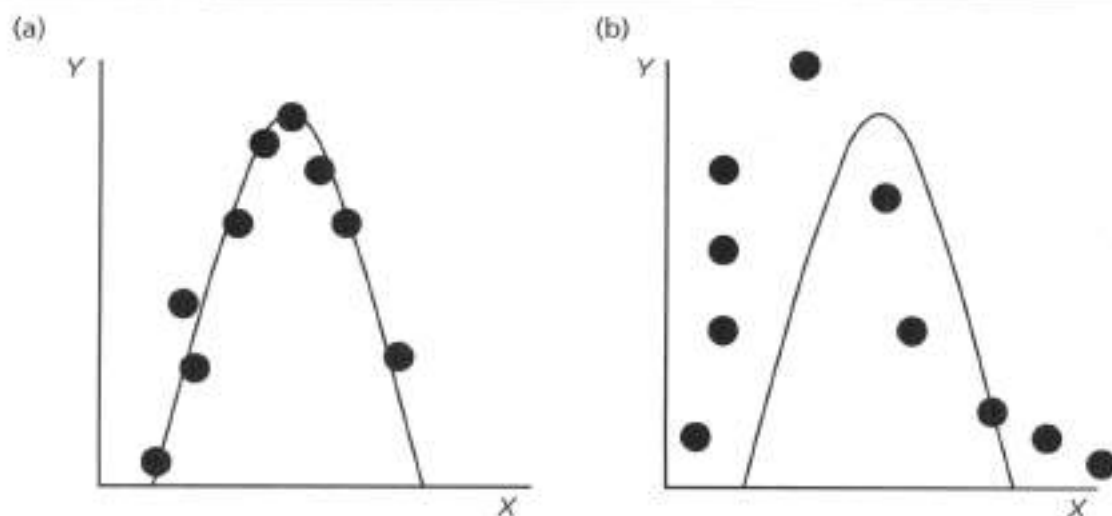
$$R^2 = 0.983$$

Indeed, as equation 12.24 and Figure 12.15 indicate, the third-degree polynomial is a better fit than a second-degree polynomial. In fact, if we keep increasing the degree of polynomial, the resulting curve will get closer and closer to the observed points and, when we use $n - 1$ (the number of observations minus 1) degrees, there is no residual. A perfect fit! Unfortunately, such an exercise reduces our result to a mathematical tautology. In fact, it is extremely rare for researchers in the social sciences to use higher than a second-degree polynomial. Apart from the tautological reasoning, the use of a higher degree polynomial imposes undue restrictions on the data. For example, in a third-degree polynomial the predicted values of Y may increase for the initial range of the value of X , then decrease rapidly, and finally increase at a dramatic rate. In the final stage, when the

FIGURE 12.15 Prediction with a Third-degree Polynomial

coefficient of the extremely large third-degree term becomes dominant, predictions outside the sample range will give a highly inflated figure, as you can see in Figure 12.15.

Even the simpler quadratic form is not free of structural biases. Its strict mathematical structure forces a symmetric specification, which can lead to faulty predictions. The result of this bias is shown in Figure 12.16. If the dependent variable shows a symmetric

FIGURE 12.16 Symmetry Forced by a Quadratic Form

distribution, then you are safe. If the distribution is skewed, however, the predictions will be suspect. Unfortunately, this bias may not be apparent from the regression results, which will show a fairly decent fit. Yet if you plot the actual observations against the predicted values, the quadratic form will reveal its shortcomings.

Log-transformed Forms

In the segment of social science research that uses **regression analysis**, perhaps the most common functional specification after the linear form is the log-log form. To understand log-log functions, you should know what a log, or logarithm, is.

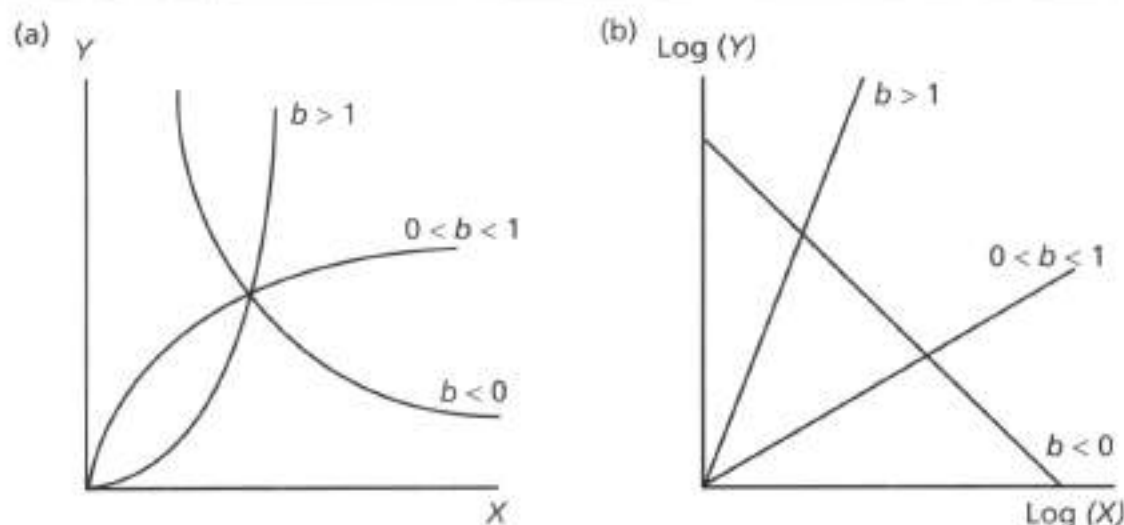
A **log transformation** expresses a number by the exponent of its base. That is, suppose we are using 10 as the base. Since 1,000 is 10^3 , we write $\log(1,000) = 3$. By following this logic, $\log(100) = 2$, $\log(10) = 1$, and because any number raised to the power of 0 is equal to 1, $\log(1) = 0$. For numbers less than 1, their log values become less than 0, or negative. However, as the logs of numbers come close to 0, they produce large negative numbers and, therefore, the log of 0 is an infinitely large negative number. So *you cannot take the log of 0*. Logs can have any base, but the most commonly used base is 10. Hence, it is called the **common log**. If the base of a log is e (a commonly used number in mathematics, equal to 2.71828), it is called a **natural log**. If you want to get the original values from the log-transformed values, you have to use an **antilog**. Thus, the common log of 10 is 1, and the antilog of 1 is 10. The common log is written as **log**, whereas the natural log is expressed as **ln**. For a better understanding of logarithms, you may try log transforming a few numbers with the help of your hand calculator.

Although you can use either 10 or e as the base, in statistics it is more common to use the e -based natural log. Therefore, in this section, we will use the expression **ln** for log-transformed variables. By log transforming the dependent and independent variables, a log-log functional form is written as

$$\ln(Y_i) = a + b \ln(X_i) + e_i \quad (12.25)$$

The log transformation of data offers us a few attractive features. First, a series may exhibit a nonlinear form, such as the ones shown in Figure 12.17 (a). If the coefficient b is greater than 1, the series would grow exponentially. Although it is hard to find examples of series with such explosive growth patterns in the realm of the social sciences, over a short period of time a series may exhibit such a pattern (for example, housing prices in the boom areas of the United States). Many series in the social sciences come close to the lines shown where $0 < b < 1$. For example, cross-national data show that as nations become affluent, their aggregate rates of growth, after periods of rapid expansion, tend to slow. Also, it is relatively easy to find examples where b is negative (< 0). For example, we know that as nations become affluent, their population numbers register steady rates of decline. In such cases log-log functions offer a better fit than linear forms. If you plot the data on a log-log form, they will show linear patterns, as in Figure 12.17 (b).

Second, the log transformation of a very large number makes it small; for instance, $\ln(85,790) \approx 11.36$. Therefore, the log transformation gives a better fitted equation for data that vary a great deal (which often is the case when the numbers are large). Note,

FIGURE 12.17 Double-Log Functional Forms

however, that if you estimate a log-log curve, the predicted values will be in log forms. To change to actual numbers, you will have to take their antilogs.

Inverse Forms

The inverse functional form expresses Y as a reciprocal (or inverse) of the independent variable X . An inverse functional form can show a positive or a negative relationship between the dependent and independent variables. You may notice that because it is an inverse relationship, the following expression will give you a negative relationship:

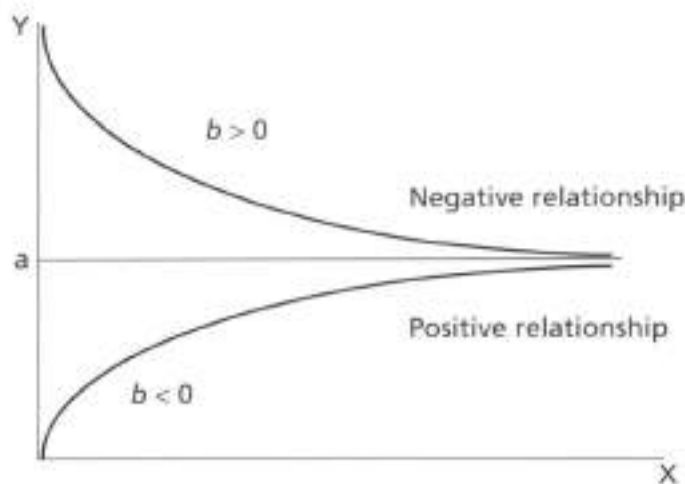
$$Y_i = a + b \frac{1}{X_i} + e_i \quad (12.26)$$

whereas

$$Y_i = a - b \frac{1}{X_i} + e_i \quad (12.27)$$

will show a positive relationship, as shown in Figure 12.18. When using an inverse form, you should remember that since you cannot divide any number by zero, there must be no zeroes in the data series for the independent variable.

You should use the inverse functional form when the value of the dependent variable falls (or rises) sharply and then approaches a certain number without actually ever being equal to it, as the value of the independent variable gets larger. Figure 12.18 illustrates the asymptotic nature of a typical inverse functional form.

FIGURE 12.18 The Inverse Form

The Problem of Irrelevant Independent Variables: Adjusted R^2

Although R^2 is widely used as a measure of goodness of fit and as a measure for choosing among alternative specifications, it contains some important drawbacks. The most important problem with R^2 is that it measures how much of the total variation is being explained by the model. Therefore, in a multiple regression model, if one keeps adding variables that are only marginally relevant, the total amount of explained variation will go up (or, at the very least, when a variable is absolutely irrelevant, the amount will remain unchanged). Thus, in a model, if we include an independent variable that adds nothing to the explanation of the dependent variable, the R^2 measure *will not go down* to reflect the inclusion of an irrelevant variable. To correct this problem, we use adjusted, or corrected, R^2 . The **adjusted R^2** is written as $\bar{R}^2$ (“ R bar squared”) and can be calculated either independently or from the calculated value of R^2 . Because it is easier to calculate it from the calculated value of R^2 , I provide the following formula:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - k} \quad (12.28)$$

where n is the number of observations, and k is the number of coefficients in the regression equation.

Note the following features in this formula:

1. There is just one independent variable—that is, where $k = 1$, $R^2 = \bar{R}^2$. Otherwise, R^2 will always be greater than $\bar{R}^2$ when there is more than one independent variable ($k > 1$). You can have only one independent variable if you run a regression equation without the intercept term—that is, if you are estimating an equation $Y = bX + e$. This is a special form of equation that makes a rather stringent assumption that the intercept

term is equal to zero. Although special reasons to formulate such a model may arise, in general you should not specify a model without the intercept term.

2. If you include an additional independent variable in the model, unlike R^2 , $\bar{R}^2$ may go up or down. If you think about it, this makes sense. Consider once again, for instance, the case of GDP per capita in Figure 12.1. The trend line appears to be linear. However, we want to see if there is a shift in trend. In that case, when we introduce the second independent variable (a dummy variable), the squared value of time, the inclusion will either add significantly to the explanation or be an irrelevant independent variable. If it adds significantly to the explanation, the value of R^2 will increase more than the offsetting effects of an increase in the number of independent variables. As a result, the value of $\bar{R}^2$ will increase from the previous estimation. In contrast, if the improvement in explanation as a result of the inclusion of an additional independent variable is less than the dampening effects of an additional variable, $\bar{R}^2$ will register a decline.

Suppose that five years ago, your city introduced tough youth crime legislation requiring mandatory punishment for gun-related crimes. You have been asked to see if the action has resulted in a change in the city's crime rate. By using a simple time-series regression ($k = 2$) on the data for the past twenty years ($n = 20$), your model explains 60 percent of the total variation ($R^2 = 0.6$), showing a gradual rising trend. From this information you calculate the adjusted R^2 as follows:

$$\bar{R}^2 = 1 - (1 - 0.6) \frac{20 - 1}{20 - 2} = 0.58 \quad (12.29)$$

Now you have included a dummy variable starting the year the legislation was passed. By reestimating your equation with two independent variables, you can attempt to provide evidence for the efficacy of the legislation. Let us suppose the new equation shows an R^2 value of 0.62, or a 2 percent improvement in your overall explanation of the city's crime rate. You can now calculate the adjusted R^2 for the new equation:

$$\bar{R}^2 = 1 - (1 - 0.62) \frac{20 - 1}{20 - 3} = 0.57 \quad (12.30)$$

As you can see, since the addition of the dummy variable did not improve the overall explanation significantly to offset the dampening effects of fraction $([n - 1]/[n - k])$, the $\bar{R}^2$ value actually went down, leading you to conclude that there is no significant statistical evidence to suggest that the passage of the legislation has had an effect on the city's crime rate.

The Significance of Coefficients Taken Together

A multiple regression model contains many independent variables, most of which are statistically significant on an individual basis. However, questions may still be raised as to their significance as a set. For this test, we test the null hypothesis that none of the explanatory variables helps explain the variation of the dependent variable around its mean (that is, $b_0 = b_1 = b_2 = \dots = b_n$). In our previous attempt to establish statistical significance

of individual coefficients, we used the t distribution. For this test, known as the joint probability test (since we are testing the collective significance of all the independent variables), we use the F distribution. We calculate the F value for the equation as

$$F_{(k-1, n-k)} = \frac{R^2}{1-R^2} \frac{n-k}{k-1} \quad (12.31)$$

The F test is closely linked to R^2 in that while testing the null hypothesis that none of the independent variables is relevant, the F test, in fact, is testing the null hypothesis that $R^2 = 0$. Therefore, for a two-variable linear equation, the null hypothesis is that the slope of the regression line is horizontal (such as in Figure 11.3 in the previous chapter).

From our example, we can calculate the F statistic as follows:

$$F_{(2-1, 5-2)} = F_{(1,3)} = \frac{0.9806}{(1-0.9806)} \frac{5-2}{2-1} = 151.64$$

The subscript for F , $(k-1, n-k)$, denotes the degrees of freedom. You will see that the F table (Appendix E) is arranged in a matrix form, where the coordinates are specified by the row and column numbers. For our test, the first number of the subscript $(k-1)$ refers to the numerator (the column), and the second number $(n-k)$ refers to the denominator (the row) of the table. By consulting the table, we can find that the critical value for $F_{(1,3)}$ at a 5 percent level of significance is 10.1. Since our F value, 151.64, is greater than 10.1, we can reject the null hypothesis and conclude that the model is indeed relevant.

Choosing the Correct Functional Form

So far we have discussed quite a few functional forms for regression models. When it comes to choosing the correct functional form, essential for good forecasting, the books on econometric theory are of little use to a researcher. However, I can offer these words of advice: *unless theory, common sense, or your experience tells you otherwise, use a linear form*. Before using any model, read theoretical literature on its behavior, plot the data to see if there is any reason to believe that a form other than a linear one is more appropriate, and draw on your or an expert's experience in choosing the functional form.

Imagine that you have been asked to project the property tax revenue for your county. After plotting the data, you detect a slight upward trend that may tempt you to go for a second-degree polynomial fit. However, you should be careful, because the trend may be the result of a recent economic boom, and its effects may be temporary. Therefore, although the estimated equation with a quadratic form is giving you better R^2 and adjusted R^2 values, your estimations may be too optimistic, particularly when the economy slows down.

FORECASTING AND ITS PROBLEMS

The purpose of estimating a regression equation is to forecast. We depend on forecasting in every facet of modern life. The stock market reacts to the projected future of the economy, revenue and expenditures are forecasted for government budgets, students

often choose their careers on the basis of the expected remuneration, orders are placed on the basis of forecasted demand, and investments are made on the basis of expected future trends in the market. Therefore, from private decisions to public policies, forecasting has become an integral part of our lives. However, you often hear such ironic comments as "It's extremely difficult to predict, especially the future." Indeed, there is a great deal of truth to this cynicism. However, since we are in the business of forecasting, let us go about it in a more systematic fashion.

A projected forecast is a quantitative estimate of the likelihood of an event's taking place in the future, on the basis of available historical data. There are two kinds of forecasts. A forecast can be either a point forecast or an interval forecast. A **point forecast** predicts a particular value for our dependent variable, which is likely to take place at a certain future point in time. An **interval forecast**, in contrast, indicates a band within which the future value is likely to lie.

Point Forecasts

We tend to prefer to offer single numbers rather than ranges, so most of our predictions turn out to be point forecasts. Interval forecasts make clear the probability factor associated with the predicted value, but point forecasts do not always do so. However, whether explicit or not, all forecasts contain an element of probability. Thus, when we forecast the government's revenue for next year to be \$1.6 trillion or for the population of a city to be 123,758, despite the apparent precision, we are predicting a probable occurrence by extrapolating from a past trend. In a sense, an economic forecaster has a certain handicap compared with a weather forecaster. We associate weather forecasting with a probabilistic outcome, even though the actual outcome is always binary—either it is going to rain or it is not. Thus, when we are told that there is a 40 percent chance of showers, we are not terribly disappointed with the forecaster when it does not rain. Yet the results of forecasts based on a regression model, with its numerical precision and its omission of the probability factor, convey a certain sense of determinism and thus can be extremely deceiving. Therefore, it is important to inquire into the factors that make a forecast good or poor.

The worth of a forecast is in its accuracy. Yet it is possible to have a good predictive model provide an inaccurate prediction, or a prediction based on a poor model turn out to be astonishingly close to reality. However, before we delve deeply into the questions of accuracy and sources of possible error, it is important to clarify certain useful terminology.

Forecasting can be either **ex post** or **ex ante**. Suppose we are predicting on the basis of a series of past data, which ends last year. We know the value of the dependent variable for this year, so we may compare the accuracy of our prediction against this known data. This is called an *ex post* prediction, in which we know with certainty the values of the dependent and independent variables. It is still a prediction because the model is providing us with values outside of the study period (which had ended the previous year). It is called *ex post*, or "after the fact," because the event has already taken place. In contrast, an *ex ante* forecast predicts values that are not yet known.

We can also distinguish between **conditional** and **unconditional** forecasts. An unconditional forecast is made when we know with certainty the values of all the

independent variables. If we are predicting on the basis of time alone, measured in years, we know the value of the independent variable in a forecast for the next year. However, if we do not know the values of the independent variables with any certainty, the forecast is called conditional. An ex post forecast is always an unconditional forecast because we know the values of the dependent and independent variables. But an ex ante forecast may be either conditional or unconditional. Clearly, if we do not know for sure the values of the independent variables in the future, this will be an example of both an ex ante and an unconditional forecast. However, like the example of time series for a future year, or when the independent variable is the past year's dependent variable (for example, when we hypothesize that how we will perform next year will depend on how we do this year), then a forecast can be both ex ante and conditional.

Interval Forecasts

An interval forecast is made when a boundary or band is provided, within which the actual value is likely to lie. For this kind of forecast, we need to use the law of large numbers again. We discussed earlier that we can test the validity of our null hypothesis by calculating how many standard deviations away the estimated parameter is from its "true value." Using the same logic, we can predict with a given level of certainty that the actual value of the predicted variable will fall within a particular range. Thus, for example, we are testing at a .05 percent level of confidence the null hypothesis that our estimated parameter b is the same as the "true" parameter β . We write this as

$$-t_{.05} \leq \frac{b - \beta}{\sigma} \leq t_{.05} \quad (12.32)$$

where $t_{.05}$ is the value of the t ratio (derived from the t table) at a .05 percent level of confidence, and σ is the standard deviation.

We reject the hypothesis if the calculated value of this ratio is greater than or equal to, or less than or equal to, the critical value of the t ratio. Now it is easy to see that we can use the same logic to state that

$$b - \sigma t_{.05} \leq \beta < b + \sigma t_{.05} \quad (12.33)$$

The preceding equation states that we are 95 percent confident that the actual value of the parameter β will be between $b \pm \sigma t_{.05}$. Therefore, we can use the same technique to derive the band within which we predict with a certain degree of confidence what the future value of our dependent variable will be. This is given by

$$\hat{Y} - \sigma_f t_{.05} \leq Y_f \leq \hat{Y} + \sigma_f t_{.05} \quad (12.34)$$

where $\hat{Y}$ is the predicted value of Y , σ_f is the standard deviation associated with forecasting, and Y_f is the future value of Y .



A CASE IN POINT

Who Wins Olympic Medals? Beijing 2008

Problem

Every four years the nations of the world get together and participate in the Summer Olympic Games. As nations compete as teams, a game quickly transforms into a platform for showing patriotic pride. Some countries' governments embrace the games as part of their national policies. At the end of the games, when the medals won by each country are tallied, there is a great disparity; some countries win a lot, while many never win anything. You may wonder why some countries do well in sporting events while others do not. We can explore this question with the help of regression analysis.

Hypotheses

Being engaged in sports requires time and money. If you have visited Olympic training facilities, you understand not just the dedication of the individual athletes but also the investment a society makes in making their young men and women competitive on the world stage. We can hypothesize that as a nation becomes more affluent, it is better able to divert resources into recreational activities. We hypothesize that per capita gross national income (GNI) will have a positive coefficient.

Winning medals depends not only on the amount of money a country can afford to invest, it also depends a great deal on the total number of athletes in the pool. Therefore, it is easy to see that more populous nations, such as China and the United States, will have a decided advantage over tiny nations like Luxembourg and Brunei. We hypothesize that population will have a positive correlation.

Apart from the size of the population and the relative affluence of a country, its probability of winning also depends on its government's active involvement in sport. During the Cold War, the communist bloc countries saw success in sports as the most visible vehicle for demonstrating their competitive edge over their more affluent and technologically advanced adversaries in the West. The tradition has continued even after the fall of the Berlin Wall. Most of the former communist nations pursue sports as a demonstration of national pride and see them as a vehicle for cementing national identity. Therefore, the third variable we can introduce in our model is a dummy variable equal to 1 for a current or former communist country and 0 otherwise. We hypothesize that the coefficient of this variable will show a positive correlation.

In our effort to explain the winning of medals, we can also introduce a fourth variable. If you examine the list of countries that win medals, you will notice the conspicuous absence of certain nations: some of the conservative Islamic nations of the Middle East. This is because their conservative tradition does not allow women to compete in sports; even men wearing immodest athletic outfits are frowned upon. As a result, these nations do not compete very well on the world stage. Once again, to account for this qualitative factor, we must use a dummy variable: these countries will have a value of 1, and all others will have a value of 0. We hypothesize that the coefficient of this variable will show a negative correlation.

Data Collection

One of the best sources of data on cross-national incomes and populations is the World Bank. We will use World Bank data, being mindful of the fact that the data set does not contain information

on some countries. For instance, because of the war, it does not post its estimate of 2008 per capita GNI for Iraq. Similarly, it does not show a population figure for Afghanistan. However, the Central Intelligence Agency's *World Factbook* gives an estimate. I have included the CIA's population data for Afghanistan. In the case of other countries for which there were no available data, I eliminated them from the sample, which consists of 136 nations.

Estimation

The hypothesized equation is written as

$$\text{Number of medals} = f(\text{Per capita GNI, Population, Communist, Middle East})$$

While estimating the equation, I came across two interrelated problems. There are wide variations in the variables of medal count, per capita GNI, and population. Therefore, I decided to log transform these variables. However, since the log of 0 is undefined, I added 1 to the series of medals and then log transformed them. This gave me the log value of 0 for those countries that failed to win a single medal.

Estimation Results

The regression model yielded the following results:

$$\begin{aligned} \ln(\text{Medals}) &= -3.41 + 0.433\ln(\text{GNI}) + 0.334\ln(\text{Population}) + 1.12(\text{Communist}) - 0.675(\text{Middle East}) \\ t &= (-8.9)^* \quad (6.7)^* \quad (10.7)^* \quad (7.23)^* \quad (-3.18)^* \\ \beta &= (0.382)(0.597)(0.414)(0.182) \\ R^2 &= 0.612 \\ \text{Adjusted } R^2 &= 0.600 \\ F_{(4,128)} &= 50.54^* \end{aligned}$$

*Significant at the 99% level

Discussion

The estimated equation explains more than 61% of the medal winnings. All the null hypotheses are rejected at the 99% level of confidence. The model clearly indicates the importance of income, population, governments' national sports policies, and culture.

Discussion Points

1. You can find the data used in this equation at www.rohan.sdsu.edu/faculty/dgupta. Calculate the predicted medal winning for each country from this equation.
2. Which countries did better than expected in the 2008 Beijing Olympics, and which did worse? Is there a pattern? Explain.
3. Can you think of other variables that can be used to make this equation even more accurate?

This formula would provide us with the necessary interval for our predicted value if we knew the value of σ_e . Since we do not know this value, we estimate it. To estimate the errors associated with prediction (S_f^2), we need to adjust the standard error of estimation (SE^2) derived in equation 12.34 with the following factor:

$$S_f = \sqrt{S_f^2} = \sqrt{SE^2 \left[1 + \frac{1}{n} + \frac{(X_{t+1} - \bar{X})^2}{\sum (X_i - \bar{X})^2} \right]} \quad (12.35)$$

where X_t is the terminal period of the study. If our study period includes 25 years, then X_t is 25, and when we predict for the next year, X_{t+1} is 26. Therefore, using this measure, we can rewrite equation 12.35 as

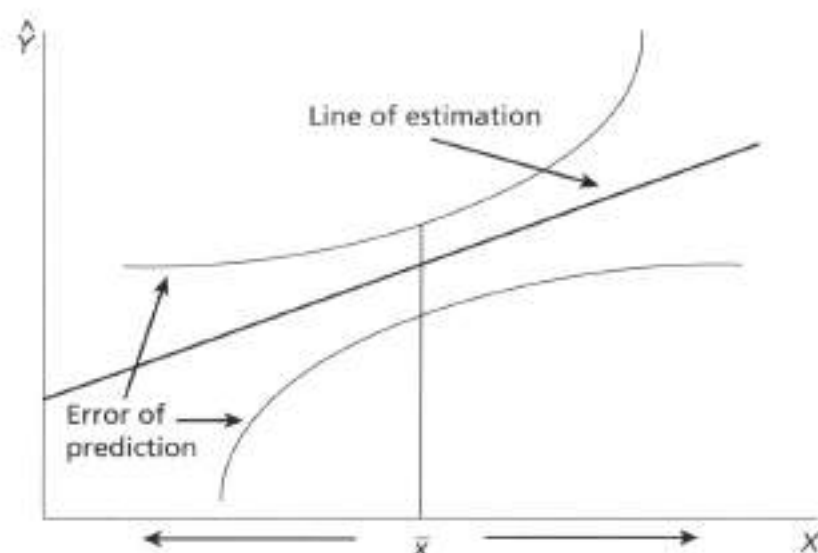
$$\hat{Y} - S_f t_{.05} \leq Y_f \leq \hat{Y} + S_f t_{.05} \quad (12.36)$$

It is interesting to note from the formation of equation 12.35 that because of the expression $(X_{t+1} - \bar{X})^2$, as the prediction moves away from the mean ($\bar{X}$), the error of prediction flares out in an exponential manner. In other words, the further into the future we want to predict, the greater our chance of committing an error and, therefore, the larger the band of prediction interval. This is shown in Figure 12.19.

Looking back at our first example of forecasting population for a small city, we had already calculated the values of S (which is 1.9) and $\sum (X_i - \bar{X})^2$ (which is 10). Hence, for us, the standard error of prediction is

$$S_f = \sqrt{1.9 \left[1 + \frac{1}{5} + \frac{(6 - 3)^2}{10} \right]} = 2.0$$

FIGURE 12.19 Errors of Prediction



This error of prediction is distributed with $n - 2$ degrees of freedom. Since we had five observations, the relevant t value is 3.182. We can derive the prediction for the sixth year from our estimated equation, which is 20.3. Therefore, the interval of prediction is given by

$$20.3 - 2 \times 3.182 \leq \hat{Y}_t \leq 20.3 + 2 \times 3.182$$

or

$$13.94 \leq \hat{Y} \leq 26.66$$

We are 95 percent confident that the future population for the sixth year will fall between the band 13.94 and 26.66. To check out the flaring effect of the prediction errors, you may calculate it for a few more years in the future.

EXPLAINING THE PRESENT WITH THE PAST: LAGGED DEPENDENT VARIABLES

In the real world, once an incident has taken place, its effects linger for a considerable amount of time. We can argue that the level of crime a city is presently experiencing will be largely dependent on the extent of criminal activities last year. This kind of a lingering relation is particularly true when the relation between the dependent and independent variables is shaped by long-term factors such as culture, poverty, or education. Therefore, it may be true that for a forecast of a city's future crime rate, we need to look at past records of criminal activities. In econometric analysis, this is accomplished by running a regression treating the past year's observation as the independent variable for this year. Thus,

$$Y_t = a + bY_{t-1} + e_t \quad (12.37)$$

In technical terms, this use of the past year's data as the independent variable is called a **lag dependent variable** or an **autoregressive model**. An example may clarify this model for you. Suppose we are trying to forecast the number of violent crimes for future years. We have the data presented in Table 12.12.

In Table 12.12 the yearly crime statistics for the city (Y_t) are presented under the heading "Current year's crime rate." In the column labeled "Previous year's crime rate," the data are lagged by one period (Y_{t-1}). Using this information in equation 12.38, we estimate that the model for the lagged dependent variable will be

$$\begin{aligned} Y_t &= 2.192 + 0.981Y_{t-1} \\ t &= (1.26) \quad (12.53)^* \\ R^2 &= 0.929 \\ \text{Adjusted } R^2 &= 0.923 \\ n &= 14 \\ F_{(1,12)} &= 156.83^* \end{aligned} \quad (12.38)$$

*Significant at the .05 level

TABLE 12.12 **Yearly Crime Statistics**

Year	Current year's crime rate (Y_t)	Previous year's crime rate (Y_{t-1})
1	10	—
2	13	10
3	13	13
4	16	13
5	15	16
6	18	15
7	19	18
8	25	19
9	24	25
10	26	24
11	27	26
12	29	27
13	30	29
14	31	30
15	35	31

You may notice that as a result of lagging, we are missing the first year's data. Hence, the number of observations has become 14 instead of 15. You may also notice that the results of this equation can be used to forecast the future year's crime rate. For instance, for the sixteenth year, we need to plug in the data for the fifteenth year in the equation. That is, the estimated figure for the sixteenth year is

$$\hat{Y}_{t+1 \text{ (6th year)}} = 2.192 + (0.981 \times 35) = 36.52$$

This equation can be used ad infinitum to generate future forecasts. For the seven-teenth year, we need to plug in the forecast for the sixteenth year; for the eighteenth year, we need the forecast for the seventeenth year, and so on. As we discussed in the previous section, the further we move into the future years (and therefore away from the mean of the actual series), the progressively larger will be the errors of estimate.

The use of a lagged dependent variable in a regression equation carries some rather important implicit assumptions. Be aware that by accepting the previous year's observation as the independent variable, you are assuming that all the external forces that shaped last year's crime rate will continue to have the same impact in determin-ing future values.

Forecasting by Curve Fitting: A Step-by-Step Approach

- *Review the literature.* Understand the underlying relationship linking the depen-dent variable with the independent variables. If there is no literature available, prepare the appropriate arguments for your postulated hypothesis.

- *Collect data.* Collect data arranged in a series. Make sure there are no missing or otherwise “contaminated” data.
- *Plot the data.* A scatter plot will show you the nature of the underlying trend.
- *Formulate a projection model.* Decide on the appropriate model of forecasting. When in doubt, use parsimony as the guiding principle. If there is a strong quadratic or exponential trend, use an appropriate nonlinear form. If in doubt, use a linear form. If there are a number of distinct trend patterns or many fluctuations in the series, use the moving average model.
- *Estimate the model.* Use appropriate software to estimate the parameters. Present your findings with all the estimated parameter values and their corresponding t values. Report the R^2 value and the F statistic. Then verbally explain the full implications of the estimated model. If the report is to be read by a nontechnical person, consider placing the results of your estimation in a footnote or an appendix. Make sure that the reader is fully aware of your methodology and assumptions and the data.

When using a trend model, you must remember that the trend method assumes the existence of a continuing trend and therefore cannot predict turning points. If the trend pattern changes or is expected to change, you are better off using a causal model of prediction.

KEY WORDS

- | | |
|---|--|
| Adjusted R^2 (p. 303) | Log transformation (p. 301) |
| Antilog (p. 301) | Method of least squares (p. 271) |
| Autoregressive model (p. 311) | Multiple regression (p. 272) |
| Coefficient for trend (slope) (p. 275) | Natural log (p. 301) |
| Coefficient of determination (R^2) (p. 279) | Point forecast (p. 306) |
| Common log (p. 301) | Population mean (p. 281) |
| Conditional forecasting (p. 306) | Population standard deviation (p. 281) |
| Dependent variable (p. 275) | Qualitative change (p. 288) |
| Error of estimation (p. 276) | Random (p. 281) |
| Error term (p. 277) | Random errors (p. 273) |
| Ex ante forecasting (p. 306) | Regression analysis (p. 301) |
| Ex post forecasting (p. 306) | Regression coefficients (p. 287) |
| Independent variable (p. 275) | Regression equation (p. 275) |
| Intercept (p. 275) | Simple regression (p. 271) |
| Intercept dummy variable (p. 288) | Slope dummy (p. 294) |
| Interval forecast (p. 306) | Statistically significant (p. 284) |
| Lag dependent variable (p. 311) | Student's t distribution (p. 284) |
| ln (p. 301) | Unconditional forecasting (p. 306) |
| Log (p. 301) | |

EXERCISES

1. The county health authority data show a 3 percent annual increase in the number of indigent emergency care cases. According to a recent survey of area hospitals, there were 3,500 such cases last year. Using a single-factor forecasting model, estimate the number of cases of indigent emergency care in your county for the next five years. If the county is paying \$1,500 per case for treatment, estimate the total monetary requirement during this period of time. Having made the estimate, comment on the reliability of your results.
2. A similar method can be used to estimate the number of homeless people, teenage pregnancies, illegal immigrants, and the like. As a part of your project, select a specific issue in your city or state and forecast its course for the next three years.
3. The table below presents three different distributions, all showing a quadratic structure. Explain which one (*Y*, *K*, or *Z*) will be best explained by a quadratic specification.

Dependent variables		Independent variables	
<i>Y</i>	<i>K</i>	<i>Z</i>	<i>Time</i>
2	2	2	1
4	3	10	2
7	5	15	3
10	12	8	4
8	16	4	5
3.5	10	3	6
2	2	2	7

4. Consider the following hypothetical yearly data on garbage collection for the town of Masters, Pennsylvania. Estimate the tonnage of garbage collection for 2009 and 2010. Use various models of forecasting. Which one of these models would you choose and why? Explain your choice of forecasting technique.

Collection of Garbage in Masters, Pennsylvania

Year	Garbage ($\times 1,000$ tons)
1987	1,325
1988	1,356
1989	1,386
1990	1,402
1991	1,435
1992	1,495
1993	1,550
1994	1,597
1995	1,645
1996	1,701
1997	1,756
1998	1,810
1999	1,899
2000	1,967
2001	1,998
2002	2,030
2003	2,780

2004	3,134
2005	3,753
2006	4,004
2007	4,139
2008	4,356

5. The crime rate in the United States has gone down steadily every year since 1998. The city of Masters is no exception. The following table shows the burglary rate per 1,000 population. Use an inverse functional form to estimate the burglary rates for 2009 and 2010.

Burglary Rate in Masters, Pennsylvania

Year	Burglary rate (per 1,000 population)
1998	2.95
1999	2.65
2000	2.42
2001	1.99
2002	1.90
2003	1.85
2004	1.81
2005	1.78
2006	1.76
2007	1.74
2008	1.74

6. The following table shows per capita chicken consumption (in pounds) for the period 1994 through 2008. On the basis of this information, forecast the per capita demand for 2009, 2010, and 2011. For these years, assume that the actual numbers were 51.6, 53.0, and 53.8. Comment on the accuracy of your model.

Year	Consumption (pounds)
1994	35.6
1995	36.5
1996	36.7
1997	38.4
1998	40.5
1999	40.3
2000	41.8
2001	40.4
2002	40.7
2003	40.1
2004	42.7
2005	44.1
2006	46.7
2007	50.6
2008	50.1

7. Predict the value of Y for X equal to 16, 17, 18, 19, and 20.

X	Y
1	35
2	28
3	25
4	20
5	18
6	16
7	15
8	16
9	21
10	27
11	32
12	35
13	39
14	45
15	53

8. Collect any time-series data (for example, government debt as a percentage of GDP, population, sales tax revenue), and with the help of a regression model, project it for the next five years. Estimate the errors of projection for each year's projection and calculate the confidence intervals.
9. The passage of the Property Tax Limitations Act (Proposition 13) in California in 1978 caused a dramatic shift in local government finances for the state. The following are the property tax rates per \$100 of assessed valuation for the city of San Diego. Estimate the tax rate for 1991 and 1992. Also estimate the intervals for those two years.

Year	Tax rate	Year	Tax rate
1971	1.959	1981	0.0860
1972	1.809	1982	0.0195
1973	1.774	1983	0.0170
1974	1.753	1984	0.0160
1975	1.753	1985	0.0147
1976	1.733	1986	0.0130
1977	1.548	1987	0.0117
1978	1.357	1988	0.0112
1979	0.131	1989	0.0103
1980	0.088	1990	0.0099

10. After reviewing the U.S. Government Accountability Office report on poverty in Appendix A, use the U.S. Census Bureau data set on poverty to develop a regression model for projecting the U.S. poverty rate. Explain your choice of dependent variables and fully explain your findings.

Notes

1. Numerous books are available on regression analysis and the method of least squares. For one of the best explanations, see Peter Kennedy, *A Guide to Econometrics*, 2nd ed. (Cambridge, Mass.: MIT Press, 1985).
2. The method of least squares has a long history, with the names of some of the most illustrious mathematicians attached to its development. This method was first proposed in 1806 by Adrien-Marie Legendre. Shortly afterward, Pierre-Simon Laplace and Carl Friedrich Gauss justified its use and demonstrated some of its useful properties. In 1812 Laplace offered proof that every unbiased linear estimator is asymptotically normal when the number of observations tends to infinity. Further, Laplace demonstrated that for the least squares estimators, the asymptotic variance is minimal. In a series of articles (1821–1823) Gauss showed that among all unbiased linear estimates, the least squares estimators minimize the mean square deviations between the true value and the estimated value. Most important, Gauss established that this relationship holds for any distribution of the errors and for any sample size. Later corroboration of his findings came through the work of Andrey Markov (1912). Subsequent developments in the least squares method were the work of mathematicians such as Alexander Aitkens and Ronald Aylmer Fisher from the late 1920s through the 1940s. This method forms the basic building block of the field of econometrics.
3. If you are not well versed in the algebra of a straight-line equation, you may engage in the following exercise. Consider two sets of equations: (a) $Y = 2 + 0.5X$ and (b) $Y = 2 + 2.5X$. By substituting values of 0, 1, and 2 for X , you can see the implications of the intercept and trend coefficient. (Y is equal to 2, 2.5, and 3 for [a] and to 2, 4.5, and 7 for [b]).
4. For those of you who are interested in the logic of derivation of the formulas for a and b , they can be derived using differential calculus. Thus, we are trying to minimize

$$S = \sum_{i=1}^n (e_i)^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2$$

Therefore, the sum of squares is minimized with respect to the values of a and b . Hence, we have

$$\frac{\partial S}{\partial a} = -2 \sum (Y_i - a - bX_i) \quad (12.3a)$$

$$\frac{\partial S}{\partial b} = -2 \sum X_i (Y_i - a - bX_i) \quad (12.3b)$$

By setting equations 12.3a and 12.3b equal to 0, we can get two equations that can be solved for two unknowns, a and b . By solving these two equations, we get equations 12.4 and 12.5 for estimating the two parameters.

5. The squared sum of explained variations is also known as the error sum of squares, and the squared sum of total variation is also referred to as the total sum of squares.
6. Edward Ames and Stanley Reiter, "Distributions of Correlation Coefficients in Economic Time Series," *Journal of American Statistical Association* 56 (1961): 637–656.
7. Once again, note that we are using two different sets of symbols for mean and standard deviation. They are μ and $\bar{X}$ for mean and σ and s for standard deviation. This is because we want to distinguish between the population, or the "true," mean and standard deviation and the sample, or "observed," mean and standard deviation. The true values have only conceptual validity, because we can never observe them. Yet in statistical theory, theorems such as the one

shown by equation 12.8 are valid only when we have the true values. Therefore, to get around the problem, we use the estimated or sample values with some necessary modifications.

8. For a history of the case, see William Kistner, "Justice for Sale?" 2009, <http://americanradio.works.publicradio.org/features/judges/>. The entire Supreme Court verdict is available at www.supremecourtus.gov/opinions/08pdf/08-22.pdf.
9. Madhavi McCall, "Buying Justice in Texas: The Influence of Campaign Contributions on the Voting Behavior of Texas Supreme Court Justices," *American Review of Politics* 22 (Fall 2001): 349–373. Also see Madhavi McCall, "Court Decision Making in Police Brutality Cases, 1990–2000," *American Politics Research*, 33 (2005): 56–80; Madhavi McCall, "The Politics of Judicial Elections: The Influence of Campaign Contributions on the Voting Pattern of Texas Supreme Court Justices, 1994–1997," *Politics and Policy* 31 (2003): 314–347.
10. The theoretical justification for dividing the sum of squared errors by $n - 2$ lies in the fact that there are n data points in the estimation process, but the estimation of the intercept and the slope introduces two constraints on the data. This leaves $n - 2$ unconstrained observations with which to estimate the errors of estimate. Hence, the number $n - 2$ is referred to as the number of degrees of freedom.
11. You may notice that here we have estimated the time trend by using the actual values of the years (such as 1991) as the independent variable. For ease of computation, you may also use numbers 1, 2, 3, and so forth for the actual values of the year. This will not change the basic relationship between the dependent and the independent variables; only the values of the estimated coefficients will change. On a computer, try estimating the same equation both ways, and then calculate the predicted or estimated values for the dependent variable. They will not change.
12. Using the two equations, you can compute the values of the dependent variables and examine the relative efficiency of the two models.
13. To see how these equations specify a quadratic relationship, consider the following two equations:

$$(a) Y = 10.0 - 2.0X + 0.5X^2$$

and

$$(b) Y = 2.0 + 2.0X - 0.5X^2$$

If you start substituting the values 0, 1, 2, 3, and so on, for X (and squaring them for X^2) and then calculating and plotting the resulting values of Y , you will see the shapes of quadratic curves.

13

The Elements of Strategic Thinking: Decision Tree and Game Theory

In March 1976 President Gerald Ford had a problem. Epidemiologists were concerned about a virulent form of the influenza virus, swine flu, which had broken out at Fort Dix in New Jersey. Experts feared that this new strain would hit the larger population of the United States in the fall, during flu season. Many scientists suggested that the flu was related to a strain that in 1918 and 1919 had caused a worldwide epidemic and taken twenty million lives.

Quickly mutating flu viruses are nightmares for health care professionals. To counter the threat of influenza, manufacturers reproduce strains of the viruses in laboratories, incubate them in eggs, and turn them into vaccines. When injected into humans, these harmless viruses become part of the shield that protects the body from natural viruses. But viruses play hide and seek with researchers by mutating, making them impervious to inoculation. To compound the problem, the protective virus must be reproduced *before* the actual infestation takes place. If people are already infected, there is not much a doctor can do. Therefore, President Ford needed to decide whether to do nothing or to start a massive inoculation program for the entire population, especially people who fell in the high-risk category (for example, older people and those with lung ailments or other chronic health problems). If he chose to do nothing, and the worst fears of the experts were realized, the nation would face a public health catastrophe. But if the experts were wrong, the president would save a great deal of money. Massive inoculation programs carry high price tags, along with the risk that some people may die from complications related to the injection.

On the other hand, if the swine flu virus infiltrated the general population after the president had assumed the associated risks and ordered a nationwide inoculation program, the president would be admired for his prudence by a grateful public, and this gratitude would translate into a considerable amount of political goodwill for

the president. However, if the threat failed to materialize, the president's policy would likely be widely ridiculed for wasting valuable public resources. Place yourself in the president's position. What would you have done? At the root of President Ford's problem was uncertainty. He did not know the future course of the virus. It is natural that a decision maker will frequently face uncertainty. So far, our discussion of how quantitative techniques aid the decision process has not addressed the question of uncertainty. Let us now see how introducing this quirk into the process affects our decisions.

GETTING A GRIP ON UNCERTAINTY

Making decisions about an uncertain future is basic to human existence. Without the benefit of hindsight, we cannot expect our decisions to be right on target. In fact, we will never have perfect or complete information about a future event. If we did, every public policy would be a success, and in the end there would be no distinction among the past, present, and future. Since the real world does not operate in this way, we must proceed strategically and establish a logical process for viewing uncertainty from an analytical perspective.

One way of looking at the unpredictable future is to follow the early work of the eminent economist Frank Knight.¹ In his seminal work Knight distinguishes **risk** from **uncertainty**. Risk is present when one can calculate the probability of a future outcome, and uncertainty is present when one cannot. When you buy car insurance, agencies use your age, education, and driving record to estimate the actuarial risk of covering the costs of probable accidents. However, when you encounter a unique event for which there is no probability estimate, you face an uncertain situation.² If you are judging the probability of an election result, a legislative outcome in Congress, or a future draft choice by a professional sports franchise in your town, you encounter uncertainty. As suggested by our discussion of probability in chapter 7, I can assert that for a future event, we can use the measures of objective probability to estimate risk, whereas for uncertainty, we must depend on subjective estimates.

You may ask, how does the concept of risk differ from that of uncertainty? The fundamental difference between the two rests on the availability of substantive knowledge. Simply put, risky events are not unique; they take place often enough for a researcher to recognize their pattern. Insurance companies, having covered hundreds of thousands of motorists, can calculate the risk of insuring an individual driver. For example, there is a pattern to how most twenty-one-year-old men drive, and insurers take their past behavior into account when setting rates. In contrast, if you want to know if I will be involved in an accident on my way home, you are dealing with uncertainty. This is because, as you may recall from our earlier discussion of probability (see chapter 7), the outcome of any single event is unpredictable; results are predictable only over a number of tries. That is, you cannot say for sure whether you will flip heads in a single coin toss, but you can predict a 50 percent outcome over a large number of tries.

The literature on uncertainty points out that there are four general sources of uncertainty, which are listed below.³ You may notice that the first three of these depend on researchers themselves. Luckily, we can use different strategies to reduce their impact. The last source, at the core of uncertainty, is something that we simply cannot avoid and must learn to accept.

The Inadequacy of Knowledge. There are events that are not entirely unknowable but depend on the knowledge and expertise of the inquirer. If I am asked to ascertain the risk for groundwater contamination from a waste treatment plant, without substantive knowledge of the field, I may fail to come up with a reasonable answer. However, if I am a properly trained expert, I am much better equipped to calculate risk levels. To reduce uncertainty, decision makers must look for the proper personnel to evaluate the odds of uncertain outcomes. If the necessary expertise cannot be found in-house, they should look for outside consultants.

Biases of Reasoning. Personal biases can cloud the mental process through which we analyze the future. These biases may crop up from a number of different sources (see chapter 5). A decision maker must be constantly on guard against such biases.

Interdependence of Human Actions. In a society, people work in groups. As a result, the final output depends not on one individual but on a number of actors. Their joint efforts shape the ultimate outcome of a project. Traditional neoclassical economics assumes that people follow their self-interest independent of others. However, research in social psychology and other branches of the social sciences has shown that those around us fundamentally affect our behavior. In dealing with others, we can never be sure of their identities, ideologies, group affiliations, or levels of commitment.⁴ Harvey Leibenstein pointed out that any manager who assumes that employees will always work “according to the book” will soon be humbled.⁵ Hence, a proper assessment of the future must include a deep understanding of human beings as social animals.

The basic core of uncertainty. The outcome of a future event depends not only on our knowledge and effort but also on factors that are completely beyond our control. Young Back Choi explains this obscured understanding with an old Chinese fable.⁶ An old man goes into the field and finds a nice horse. “Lucky me!” he exclaims and gives it to his son. Unfortunately, his son falls off the horse and becomes lame. “This horse has brought me only misfortune,” sighs the old man. Soon war breaks out with the dangerous barbarians, and the emperor’s army takes away all the young men of the village to fight the enemy. But the army does not take the old man’s lame son. The old man is thankful for finding such a lucky horse.

DECISION MAKING AND EXPECTED PAYOFF

The question of how to make the best possible decision under uncertain conditions has come under intense scrutiny. The first analytical breakthrough was derived by John von Neumann and Oskar Morgenstern, two mathematicians from Princeton University, and was developed further over the years by a great number of mathematicians and economists. Let me explain the basic precepts of their analysis. Suppose someone offers you a choice. If you predict correctly, you may win \$2 in a coin toss or \$5 in a roll of the die. Which option do you choose? Theoretically, you should choose the coin toss because your chances of winning are $1/2 = .5$. On the other hand, with the roll of the die, your chances of calling the right number are $1/6 = .167$. The **expected payoff** in a risky situation is calculated by the following equation:

$$\text{Expected payoff} = \text{Probability of winning} \times \text{Amount of reward.}$$

That is, in our case, the expected payoffs are

$$\text{Expected payoff for coin toss} = .5 \times \$2 = \$1$$

and

$$\text{Expected payoff for roll of the die} = .167 \times \$5 = \$0.83.$$

As you can see, you will be better off betting on the coin toss—that is, unless you are a real gambler or risk taker. In that case, you may prefer the higher reward offered by the die regardless of the chances of winning. For the moment, let us assume that you are neither an excessive risk taker nor an extreme risk averter; instead, you are a risk-neutral decision maker. If you want to maximize your chances of winning in the long run, you will be better off, given uncertainty, following the law of rational decision making and choosing the option with the highest expected payoff. In fact, in every uncertain situation, including those involving gambling or card games, good players will depend primarily on the proper calculation of the odds. At blackjack tables in casinos, the house realizes its profits solely by playing the odds, which are in its favor. Although there are winners among the players, the house always cleans up the table at the end of the day.

THE DECISION TREE

We can use the insights developed by von Neumann and Morgenstern to analyze the optimal course of action. The branch of social science and applied mathematics that is dedicated to the study of decision making under uncertainty is broadly called game theory. However, when a single decision maker evaluates the choice of action, he or she uses a **decision tree**, a diagram showing the sequence of events with their corresponding probability figures.

Let us go back to the uncertain public policy problem of President Ford. Suppose that the experts at the Centers for Disease Control and Prevention estimate the probability of the deadly virus's reaching the United States to be 40 percent. Also suppose that the president's political advisers appraise the benefits of his decision on a scale of +10 through -10, with +10 being the most desirable outcome and -10 being the most undesirable. Because we do not know President Ford's reasoning process, we must hypothesize that if he decides to do nothing and the epidemic becomes real, this outcome will be the least desirable of his options (scored -10). But if he decides to take no action and the epidemic does not materialize, then he has remained calm in the face of an unjustified doomsday prediction and has saved a considerable amount of public money. On his scale this result rates +5.

On the other hand, if he decides to inoculate the population, the most desirable situation arises, from his perspective, if the threat of epidemic turns out to be real. In that case the president will be hailed as a savior—an outcome worth +10 on his rating scale. However, if the epidemic does not show up, the decision to start a mass inoculation program may turn out to be a political liability, which is assessed a -5 rating. Having put numerical values on the probabilities and payoffs, we can calculate the expected payoffs facing the president (see Table 13.1).

TABLE 13.1**Hypothetical Expected Payoff Matrix for President Ford's Response to Possible Swine Flu Epidemic**

Action (1)	Situation (2)	Probability (3)	Payoff (4)	Expected payoff (5) = (3) × (4)
Do nothing (A)	Epidemic occurs	0.4	-10	-4
	Epidemic does not occur	0.6	+5	3
Total		1.0	-5	-1
Start mass inoculation (B)	Epidemic occurs	0.4	+10	4
	Epidemic does not occur	0.6	-5	-3
Total		1.0	5	1

I have drawn the logical structure of President Ford's actions in the form of a decision tree (see Figure 13.1). The branches of a decision tree can be divided into **actions** and **outcomes**. Although actions are deliberate and reflect an individual's conscious decision, outcomes are uncertain. If you inspect Figure 13.1 closely, you will see that it contains four important elements:

1. *Decision nodes*: the points at which a decision maker must choose a possible course of action. In this particular case, President Ford has one decision node that shows two possible actions: do nothing or inoculate. I have represented this node as a square.
2. *Chance nodes*: the points showing the probable outcomes of an action. After the president chooses an action, there are two possible outcomes: the flu breaks out or it does not break out. The chance nodes are indicated with circles.
3. *Probabilities*: estimates of the possible outcomes. Probabilities can be determined either through observation in the case of objective probabilities or through the subjective assessments of experts.
4. *Payoffs*: consequent rewards and losses experienced by the decision makers. Payoffs can be estimated in monetary terms or, as in this case, in terms of some other agreed-upon scale. These payoffs can be actual or perceived profits or losses.

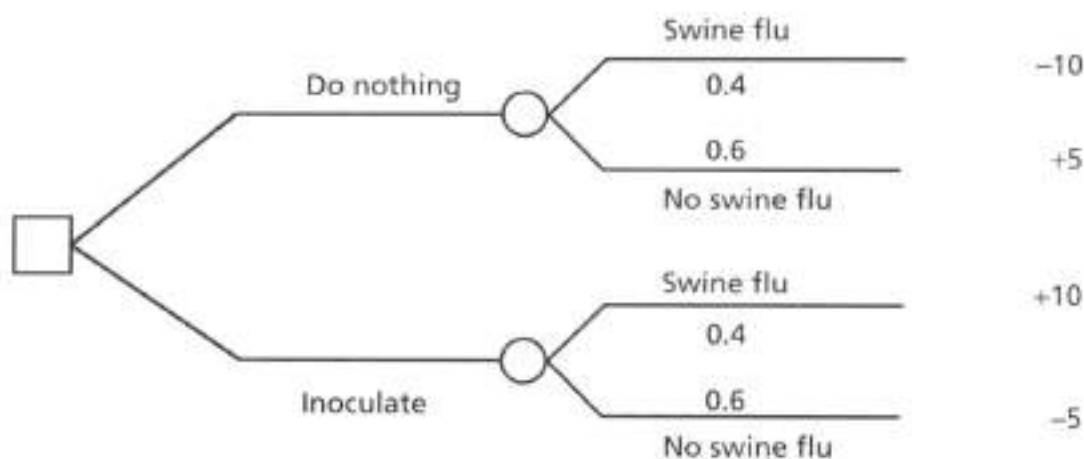
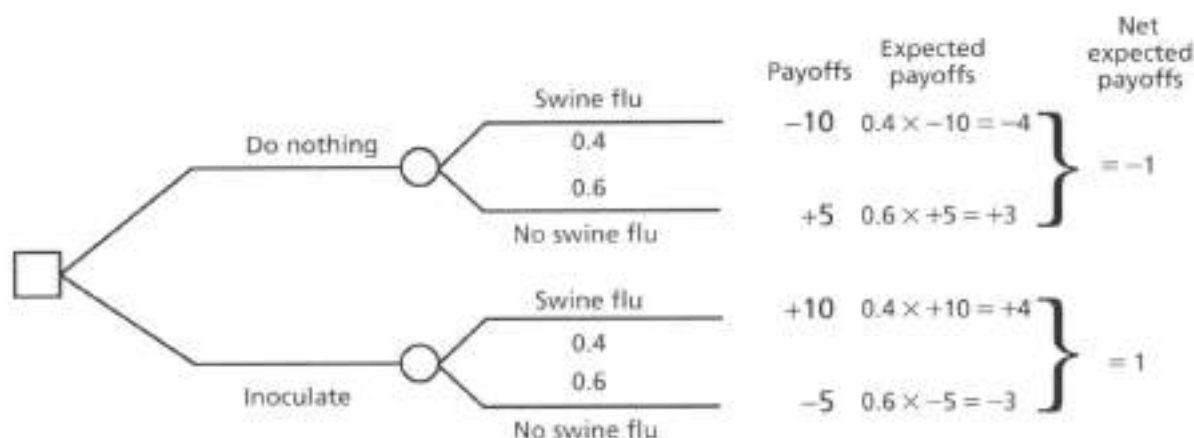
FIGURE 13.1**Hypothetical Decision Tree for Swine Flu Threat**

FIGURE 13.2 Decision Tree for Swine Flu

In Figure 13.2 we can see that, facing an uncertain situation, President Ford had a choice between an expected payoff of -1 (for the “do nothing” option) and +1 (for the “inoculate” option). Therefore, given the logical construct, he should have chosen the option to inoculate the entire population against a probable swine flu infestation. As an anecdotal postscript to this analysis, we should note that President Ford did choose to inoculate, and the virus failed to show up in the United States.⁷

Structuring a Decision Tree

From the preceding rendition of a decision tree, you can see that a tree can be a powerful tool of logical reasoning when the outcomes are shrouded in risk. This exercise allows you to see the true nature of a problem more clearly by providing you with a deeper understanding of the role of chance as it relates to the alternative courses of action. Also, the very process of identifying the alternatives, their outcomes, and their sequential interaction with their respective probabilities can reveal more about the core of the problem than you may otherwise have known.

To be an effective tool of analysis, decision trees must be as exhaustive as possible in identifying every conceivable outcome. Thus, in the above example, I have identified two options for President Ford: to do nothing and to inoculate the entire population. If there were other options, such as inoculating only those who were at the highest possible risk, or restricting the inoculations to specific geographic areas, they also should have been fully explored. This process of identifying alternatives, then, should be **collectively exhaustive**. That is, no feasible option should be deliberately left out without an explicit justification for its exclusion.

Furthermore, the options must be **mutually exclusive**. In other words, the options should be defined in ways that do not overlap. Let us say that we have defined the president’s three options as (1) do not inoculate, (2) inoculate only the high-risk population, and (3) inoculate everyone. In this case, all options are mutually exclusive, since you cannot meaningfully choose more than one option at one time. If I choose to inoculate only the high-risk segment of the population, that choice would preclude my choosing to inoculate everyone in the nation. However, suppose the president has

another option: to immediately launch a research project investigating the causes of the pandemic. The decision tree would have four branches: (1) do not inoculate, (2) inoculate only the high-risk population, (3) inoculate everyone, and (4) start a research project. As you can see, launching a research project does not exclude the other three options, and you can conduct research along with any other action. Therefore, unless your action is predicated on the findings of the research project, you should not include the project in the decision tree.

Evaluating Flood Damage Reduction

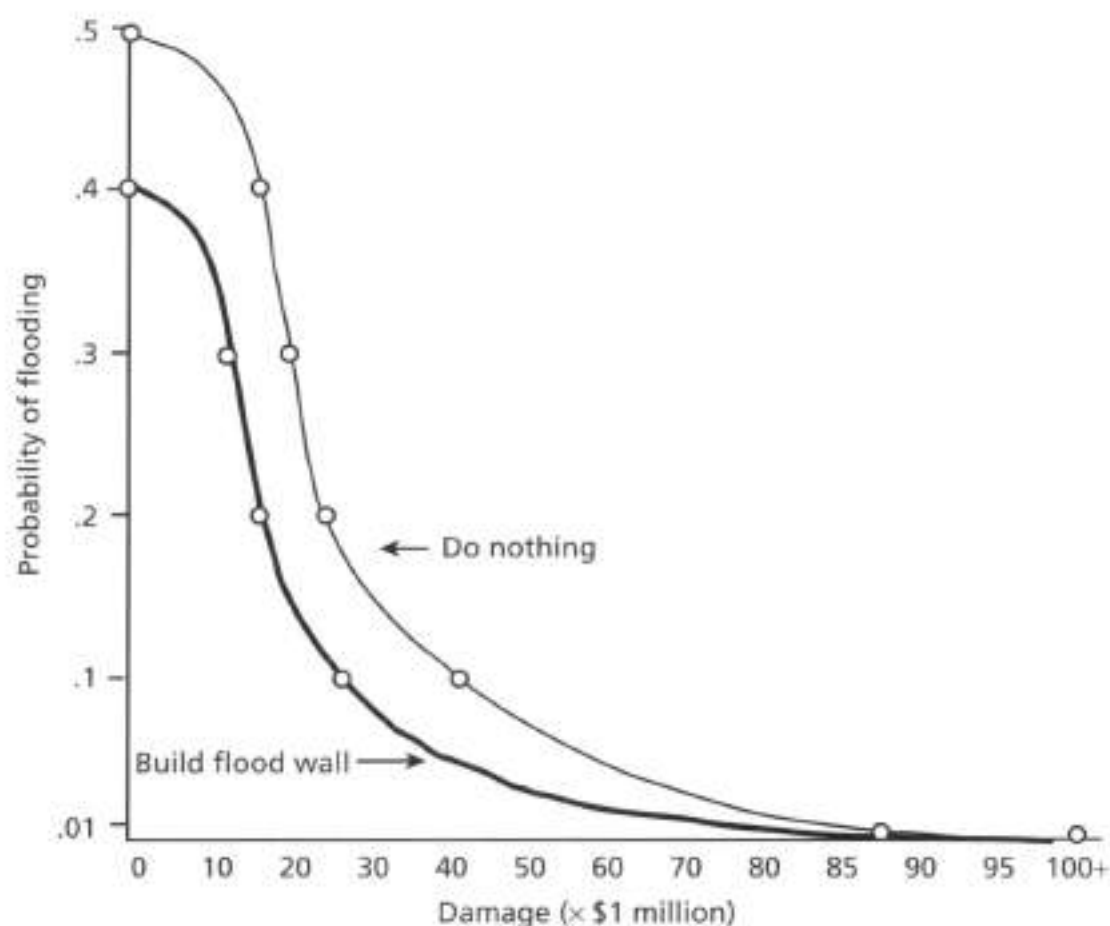
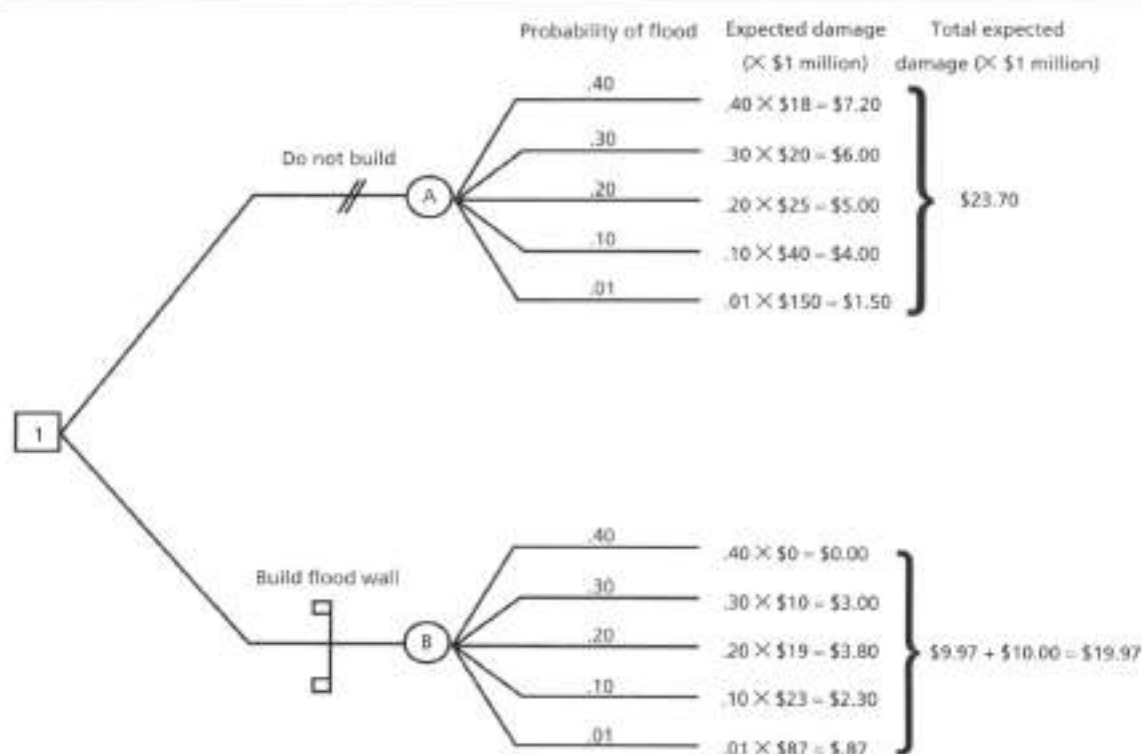
Imagine that several neighborhoods in your town have serious problems with seasonal flooding. A recent study commissioned jointly by the Federal Emergency Management Agency and the U.S. Army Corps of Engineers (COE) shows that if your town is hit by a catastrophic hundred-year flood (with a corresponding probability of 1 percent, or .01), the damage may equal \$150 million. However, the study also suggests that the construction of a flood wall may significantly reduce the damage from flooding. Table 13.2 presents the estimated damage figures corresponding to the various levels of flooding. The flooding we can expect every other year (with a probability of .5) is not likely to cause much damage. However, a more severe flood that inundates the town every five years (having a probability of .2) may cause up to \$25 million of damage. Figure 13.3 plots the data presented in Table 13.2.

As a preventive measure, the town is considering the construction of a flood wall. The COE estimates that the flood wall will cost the town \$10 million. The town council is trying to decide whether to finance the structure with a municipal bond issue. I have represented the choices open to the town council in the form of a decision tree (see Figure 13.4). To make the diagram less complicated, I have eliminated the most frequent flooding, which does negligible damage. As you can see, I have numbered the decision nodes and added letters to the chance nodes. The symbol $\square$ indicates the point in the decision process at which a fixed cost has been incurred. This symbol stands for the cost of constructing the flood wall, which we must add to the total expected flood damage.

If you look again at Table 13.2, you may be tempted to conclude that the total savings for the project is \$114 million, the total difference in estimated damage, minus

TABLE 13.2 Estimated Flood Damage with and without Flood Wall

Probability of flood	Estimated damage ($\times$ \$1 million)		Difference in estimated damage ($\times$ \$1 million)
	<i>Flood wall</i>	<i>No flood wall</i>	
.50	0	0	0
.40	0	18	18
.30	10	20	10
.20	19	25	6
.10	23	40	17
.01	87	150	63
Total	139	253	114

FIGURE 13.3 Probability of Flood Damage**FIGURE 13.4** Decision Tree for Flood Control

\$10 million, the cost of building the flood wall. However, when you take into account the respective probability measures, you will see that the bulk of the savings comes from limiting the damage of the most devastating flood, which is expected to occur every hundred years. Since the probability of such a monster flood is small (.01), the expected payoff is puny. As you can see from Fig. 13.4, the expected loss without the flood wall is \$1.5 million, compared with \$0.87 million with the flood wall. The expected savings comes to \$630,000. The total benefits of constructing a flood wall are still greater than those of not constructing it. If we go ahead with the construction, our expected damage is \$9.97 million plus the cost of construction, \$10 million, or a total of \$19.97 million. Because this number is smaller than the expected damage of \$23.7 million without the flood wall, we should go ahead with the project.

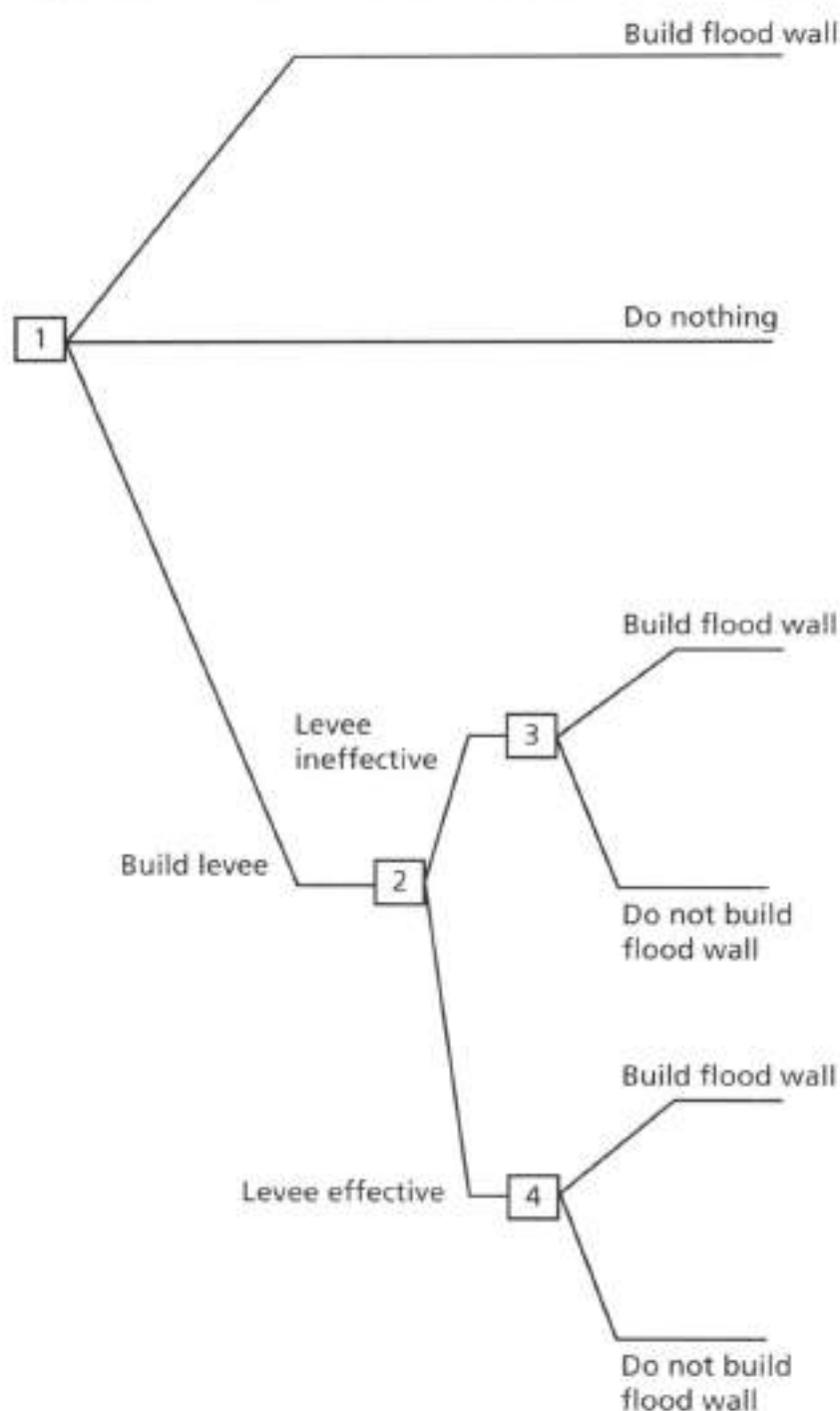
At this point you should note that since we find the “do nothing” option decidedly inferior to the construction of the flood wall, we have crossed it out with an etch mark: //. This mark shows that we have considered this option and rejected it in favor of another.

This example of a decision tree is relatively simple, since it contains only one decision node. However, a decision tree can be much more complex and include many more decision nodes. Let us extend our example to include some additional decision options.

The COE study has come up with another flood control plan. It suggests that the town council consider building, in addition to the flood wall, a small levee of locally available impervious soil. This levee, to be constructed a few miles upstream, may divert enough water to significantly reduce the town’s damages, particularly from the more frequent swells. However, there are a couple of caveats. First, the levee may turn out to be ineffective. In that case, the town will incur the same amount of damages as it would have sustained without this new construction. But if the levee is effective, the town will realize significant savings by constructing the flood wall. The COE estimates the construction cost for the levee to be \$1 million. The agency further predicts that the probability of the levee’s proving effective is 70 percent.

TABLE 13.3 **Estimated Flood Damage with Flood Wall and Levee**

Probability of flood	Estimated damage (× \$1 million)			
	<i>Flood wall</i>	<i>No flood wall</i>	<i>Effective levee with flood wall</i>	<i>Ineffective levee with flood wall</i>
.50	0.00	0.00	0.00	0.00
.40	0.00	18.00	0.00	0.00
.30	10.00	20.00	0.00	15.00
.20	19.00	25.00	5.00	20.00
.10	23.00	40.00	10.00	40.00
.01	87.00	150.00	180.00	150.00
Total expected flood damage (probability × damage)	9.97	23.70	3.80	14.00

FIGURE 13.5**Expanded Decision Tree for Flood Control with Levee Option**

The second caveat is that in the case of a catastrophic flood, the levee is likely to collapse. If it does, the damage from the flood will escalate even more than it would if the levee and flood wall had not been built (see Tab. 13.3). This caveat introduces new options. The town can still decide to do nothing, it can immediately start building a flood wall, or it can wait to see whether the earthen levee is effective. If the levee is ineffective, then it can consider constructing the flood wall. If the levee is successful, it may go ahead and build the flood control device or decide not to build it after all (see Fig. 13.5).

FIGURE 13.6

Decision Tree for Flood Control with Total Expected Costs

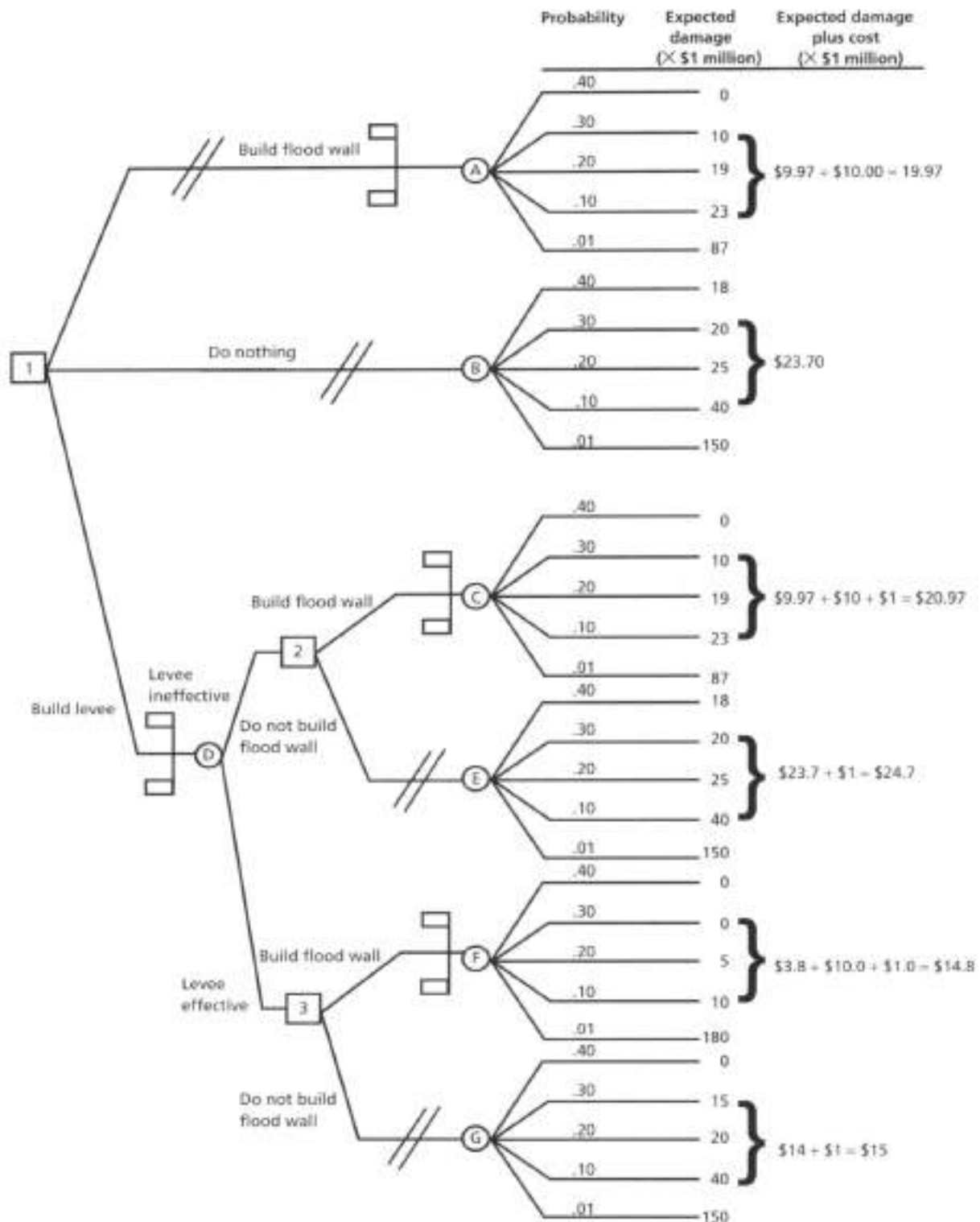


Table 13.3 presents the damage estimates for an effective levee. You may note that I have left out one option: I did not include the damage estimates for when the levee is ineffective. This is because these estimates are the same as those for no levee at all. Combining information from Table 13.2, Table 13.3, and Figure 13.5, we can build the decision tree shown in Figure 13.6.

When there are several decision nodes, the simplest way of working out a decision tree problem is *folding backward*. In this case, the first two options of doing nothing and building the flood wall must be evaluated against the option of constructing the levee first. In evaluating that option, we must consider decision nodes 2 and 3. For decision node 2, if the levee is ineffective, it stands to reason that the flood wall should be built, since the total estimated damage for the option of “do not build” (\$23.7 million + \$1 million for levee = \$24.7 million) is higher than that for the decision to construct the flood wall (\$9.97 million + construction cost of \$10 million + \$1 million = \$20.97 million).

If, on the other hand, the levee is found to be effective (at decision node 3), building the flood wall once again seems to be a better option than not building it. That is because the expected damage with an effective levee and a flood wall, \$14.8 million (\$3.8 million of expected flood damage + \$10 million for the flood wall + \$1 million for the levee), is lower than the expected cost of not building a flood wall, \$15.0 million (\$14.0 million + \$1 million).

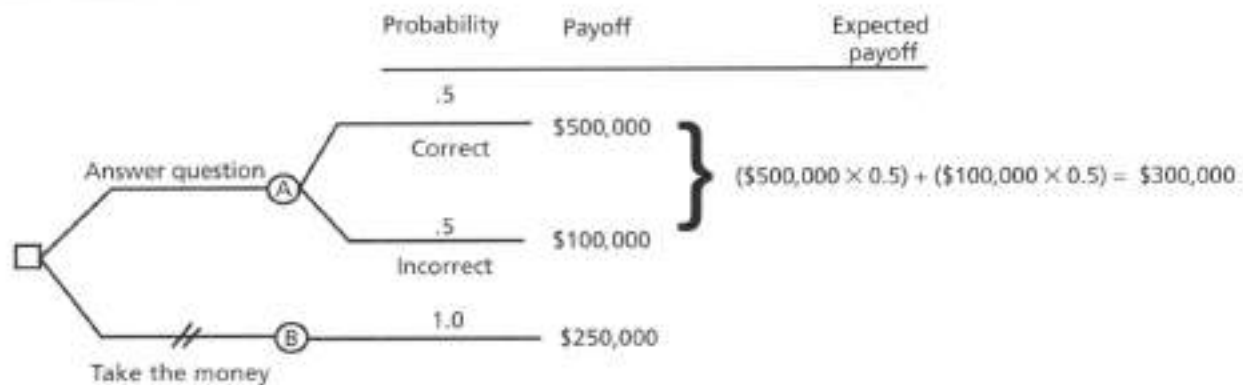
We are now in a position to evaluate the options: do nothing or immediately build the flood wall. As you can see, the estimated damage for the do-nothing option is \$23.7 million, while the damage for the immediate construction of the flood wall is expected to be \$19.97 million. Since these estimates are higher than those for building the flood wall whether or not the levee is effective, we should recommend that the town council build the levee and the flood wall. I have summarized the findings in Table 13.4.⁸

Risk Tolerance and Expected Payoff

Each of us views risk differently. I was recently watching a game show in which the contestant had two answer choices. If she gave the right answer, she would earn a

TABLE 13.4 Total Estimated Cost of Flood Wall and Levee

Probability of flood	Estimated damage (× \$1 million)					
	Ineffective levee				Effective levee	
	Flood wall	No flood wall	Flood wall	No flood wall	Flood wall	No flood wall
.50	0.00	0.00	0.00	0.00	0.00	0.00
.40	0.00	18.00	0.00	18.00	0.00	0.00
.30	10.00	20.00	10.00	20.00	0.00	15.00
.20	19.00	25.00	19.00	25.00	5.00	20.00
.10	23.00	40.00	23.00	40.00	10.00	40.00
.01	87.00	150.00	87.00	150.00	180.00	150.00
Total	9.97	23.70	9.97	23.70	3.80	14.00
estimated damage						
Cost of flood wall	10.00	0.00	10.00	0.00	10.00	0.00
Cost of levee	0.00	0.00	1.00	1.00	1.00	1.00
Total cost	19.97	23.70	20.97	24.70	14.80	15.00

FIGURE 13.7 Risk and Payoff in a Television Game Show

\$500,000 prize, doubling her current winnings of \$250,000. But if her answer was incorrect, she would receive only \$100,000. What would you do if you were in her position? Let us draw her options in the framework of a decision tree (see Figure 13.7).

If her chances of choosing the right answer are 50-50, her expected payoff is greater than that of the sure-bet option (probability = 1.0) of taking the money she has already won. But would you be a cool statistician and risk losing \$150,000? If you would, your behavior would be considered **risk neutral**. On the other hand, if you would choose the second, sure-bet option, your action would be regarded as **risk averse**; that is, you would prefer to avoid risk in decision making. In contrast, some people thrive on taking high risks. If you are only 30 percent sure that you know the correct answer, the expected returns for attempting to answer are \$220,000, \$30,000 less than the returns expected for taking the money. In this case, if you still insist on chasing the high reward of half a million dollars, you will be considered a **risk lover**.

Is it then irrational to be anything but risk neutral? There is a problem with defining human rationality as risk-neutral behavior. As I have shown in chapter 5, the so-called Allais paradox demonstrates that people's preference for risk can vary across the range of probability and reward. From a policy perspective, it may be advisable to take risks as long as you can comfortably bear the consequences of the loss. Consider, for example, our hypothetical example of constructing a levee along with a flood wall. By following the rules of mathematical expectation, we have recommended the construction of both. However, if you examine the payoff matrix carefully, you will see that such a recommendation may be shortsighted. Building both flood control devices significantly lowers the risk of more frequent flooding, but it heightens the risk for damage associated with a cataclysmic flood. A huge increase in the water level may suddenly break the levee, and the breach could exacerbate an already catastrophic incident. The question is, can you afford a small risk that you will incur a devastating loss? After all, the hundred-year flood may not wait another hundred years to show up. If it occurs, the entire town may be wiped out. Given this possibility, can you blame the authorities for taking the less expensive and less risky option of building only the flood wall? For another example of how policymakers weigh risk against expected payoffs, see "A Case in Point."



A CASE IN POINT

Risky Business: Orange County Bankruptcy

On December 6, 1994, officials of Orange County, California, one of the most affluent and fiscally conservative counties in the nation, declared bankruptcy, the result of risky investments made by county treasurer Robert Citron.⁹ When the final tallies were made, the county had lost a total of \$1.64 billion from its investment pool, which invested money for the county, its thirty-one cities, and more than 150 other, mostly single purpose, government entities.

The primary guiding principles for local government investment pools are safety, liquidity, and yield. That is, when investing public money, pools must safeguard invested money and not take unacceptable risks. Fund managers should also make sure that the municipality has enough liquid cash on hand to meet its obligations. Finally, cash managers seek to maximize return on investment (yield).

Citron was the county treasurer for more than twenty years. He was widely considered a "financial guru," who could consistently earn returns vastly superior to those realized by other municipalities throughout the state and the nation. In fact, Citron had promised that a full 35 percent of the county's 1995 revenue would come from interest income. As it turned out, Citron was paying more attention to the possibility of yield than to the reality, virtually ignoring the safety of his investments by choosing high-risk instruments called "derivatives." The problem with derivatives is that if the market follows your prediction, these securities give you extremely high returns, but if it does not, they can result in enormous losses.¹⁰

To be sure, Citron did have his critics, who claimed that he was taking unacceptable risks. However, the county, like most other local governments in California, was strapped for tax revenues because Proposition 13 (a sweeping property tax reduction measure) had put the county in a financial straightjacket since its passage in 1978. Because these investments brought in huge sums of money—without requiring that the county impose higher taxes on its highly conservative electorate—such words of caution were ignored by the politicians who supervised Citron.

TWO ACTIVE PLAYERS: GAME THEORY

The decision tree is a schematic representation of decision making under uncertainty. It falls under the more general area of study called **game theory**. Game theory is a mathematical technique to evaluate the strategic interaction between two or more opponents on the basis of expectations about their possible moves. This technique is particularly useful when you are in a situation of active confrontation, must make strategic decisions, and are uncertain about the moves of your opponent(s). The literature on game theory is varied and rich. The beauty of this technique is that it can be used to explain strategies in a variety of scenarios, from chess games to corporate takeovers, from labor union bargaining to the negotiation of international treaties. Also, because it lends itself to mathematical modeling, literature about it varies from simple to extremely complex,

In the end, Robert Citron, the “financial wizard,” was convicted of fraud. Most interestingly, during his trial, a neuropsychologist testified in his defense that she found extensive damage in Citron’s frontal brain, which “allows you to think, analyze information, and be conscious.”¹¹ Echoing her findings, a clinical psychologist compared him to “an empty bottle put out into the water.” The experts claimed that Citron had suffered from brain damage throughout his life.

Derivatives attempt to manipulate the future risks within a complex mathematical model. As a result, many who invest in these products do not fully understand the risks involved and the assumptions that go into the construction of these models. As they pushed Orange County to bankruptcy in 1994, a year later, the British banking giant Barings fell into the same trap when an errant trader in one of its overseas branches gambled on these extremely high risk products. The mounting losses brought down Britain’s oldest merchant bank, which had financed the Napoleonic wars, the Louisiana Purchase, and the construction of the Erie Canal. Investment in derivatives, particularly by hedge funds, which often operate outside banking regulations, was also blamed for the financial debacle of 2008. The biggest moral of the saga of the derivatives for the money managers is that they must not invest in anything they do not fully understand. Since the implications of investing in these exceedingly lucrative instruments are extremely dire, and so is the temptation for fund managers to seek them out for quick profits, it appears that the government and regulatory agencies may soon step in order to mitigate their devastating effects in a globalized economy.

Discussion Points

1. Does it always pay to be risk neutral when dealing with an uncertain future? Is it advisable to be a risk lover in certain situations? When?
2. From the standpoint of risk management, why did Orange County face a fiscal disaster?

from eminently practical to highly esoteric.¹² The rules of strategic decision making have been researched by psychologists (to study fallacies about uncertainty), international relations scholars (to predict the strategic moves of nations), and students of corporate strategies (to explore the ways in which corporations interact).¹³ Students of negotiating and bargaining have tried to come up with rules for conflict resolution.¹⁴ These diverse inquiries have significantly advanced our knowledge about strategic decision making and sometimes have inflated expectations beyond what these techniques can deliver.¹⁵

Let us start with a simple game with two players. Although decision trees are shown as diagrams, two-player games are usually written as matrices. Since game theory can explain the outcome of a confrontation between two or more parties, it is useful in analyzing policies of engagement. Let us assume that your city is engaged in a bitter

TABLE 13.5

Payoff Matrix for Mayor and Garbage Collectors' Union

Union's position	Mayor's position	
	Accept arbitration	Hang tough
Accept arbitration	3, 2	0, 5
Strike	4, 0	1, 1

labor dispute. The garbage collectors' union is threatening to go on strike during the busy Christmas season. The city has been in negotiation with the union for some time, but the parties remain far apart in their positions. The union believes that the conservative mayor wants to weaken organized labor and has little faith in her commitment to an acceptable negotiated settlement. The mayor, on the other hand, is up for reelection in February. She believes that the actions of organized labor are politically motivated and designed to undermine her candidacy. The state mediators, unable to end the impasse, suggest binding arbitration, in which case an independent board reviews the situation and proposes a compromise binding on both sides. From the perspective of the union, there are two options: to accept arbitration and risk a less than acceptable contract or to go on strike. Similarly, the mayor has two options. She can accept arbitration and run the risk of approving a contract she believes might be fiscally irresponsible for the city, or she can hang tough and not negotiate with the garbage collectors' union. We can analyze the city's dilemma with the help of game theory. Table 13.5 shows the payoff matrix for the two parties.

The table presents the perceived payoffs for the two parties on a scale ranging from 0 (least desirable) to 5 (most desirable). As you can see from this table, the union has two alternatives. It can accept arbitration or go on strike. If it negotiates with the mayor for an arbitration board of its liking, its preference ranking is 3. This, however, is not the best option for the mayor, since she would prefer to take a tough stance against organized labor. In any case, she assigns this option a lukewarm 2.

From the union's perspective, the worst scenario is one in which the mayor refuses to negotiate and the union is forced to accept arbitration. This combination of strategies is valued as 0 by the union and 5 by the mayor.

The union may opt for a work stoppage. If this action brings a reluctant mayor to the negotiating table, it is the most preferred option for the union (4) and least preferred (0) for the mayor. Finally, if the union strikes and the mayor hangs tough in a high-risk game, both parties assign 1 to this option.

The rules of rational choice under uncertain conditions dictate that each player follow a **minimax strategy**, in which each chooses the strategy that minimizes his or her maximum loss or regret. From the perspective of the union, if it chooses to accept arbitration, its maximum loss is to be pinned to the mat by a tough mayor. Thus, for this option, its maximum regret position is 0. However, for the strike option, its worst possible outcome is 1. Hence, from the standpoint of the union, deciding to strike is superior to going for arbitration. Similarly, you can see that the mayor minimizes her

maximum loss when she takes a strong position. Therefore, the only possible outcome of this confrontation is a strike by the garbage collectors while the mayor takes an intransigent posture.

This outcome, in the lower right-hand corner of the matrix, is called the solution of **Nash equilibrium**, after noted mathematician John Nash. This solution identifies an outcome from which neither player can gain by unilaterally switching to another strategy.¹⁶ That is, if the mayor chooses to confront the labor union, the union does not gain by unilaterally switching to a compromise option of going for arbitration, since doing so would lower its payoff from 1 to 0. Similarly, if the union chooses to strike, the best option for the mayor is to project a firm position; otherwise, her own payoff would sink from 1 to 0.

The need to devise military strategies in the face of uncertainty has inspired much of the development of game theory. One of the textbook cases of game theory can be found in the Battle of the Bismarck Sea.¹⁷ In early 1943 the allies' counterattack on the Japanese navy was in full swing. By that time U.S. forces had gained a foothold on the island of New Guinea. The northern half of the island was controlled by the Japanese, while the allies controlled the southern part. Intelligence reports produced for Douglas MacArthur, the supreme commander of the allied forces, indicated that the Japanese were sending huge reinforcements for a major counterattack on the allies. General MacArthur ordered General George C. Kenney, commander of the allied air force, to challenge the Japanese convoy and inflict the maximum possible damage. Figure 13.8 presents General Kenney's dilemma.

As you can see, the commander of the Japanese convoy had two options: he could take the northern route or the southern route. Therefore, Kenney also had the corresponding options of sending the bulk of his interceptor forces to the north or directing them to the south. His objective was to deliver the maximum possible

FIGURE 13.8 Strategies in the Battle of the Bismarck Sea

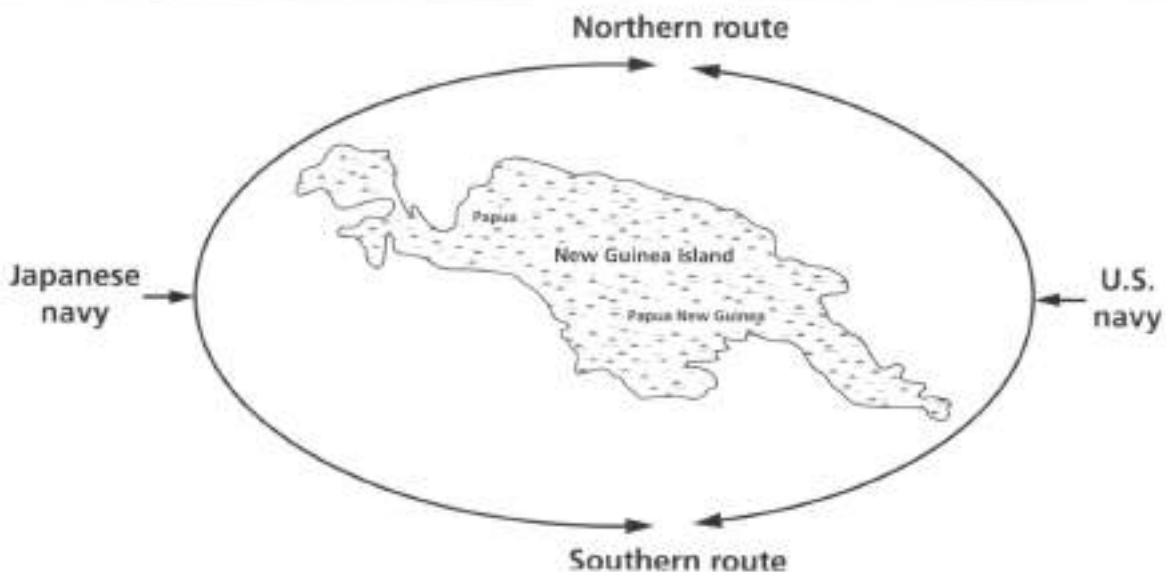
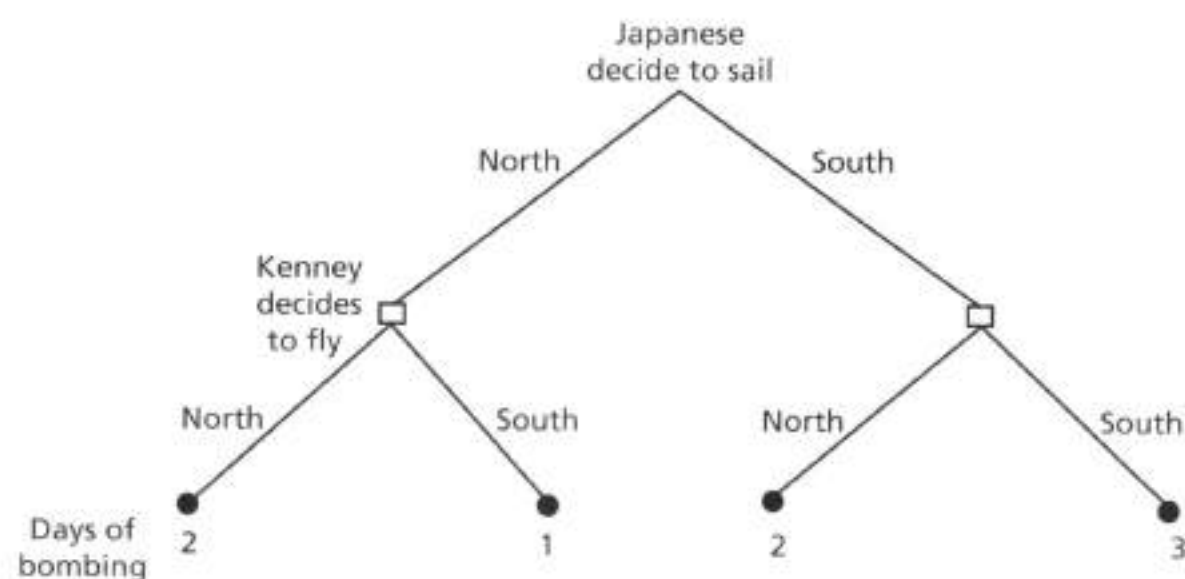


TABLE 13.6 Possible Days of Allied Bombing

Allied strategy	Japanese strategy	
	<i>Sail north</i>	<i>Sail south</i>
Search north	2 days of bombing	2 days of bombing
Search south	1 day of bombing	3 days of bombing

damage to the Japanese convoy through relentless bombing. The weather report showed an approaching storm on the north side of the island. The storm would markedly reduce visibility and, consequently, the ability of the allied planes to locate and strafe enemy ships. General Kenney measured his payoff in terms of the expected number of days the allied air force would have to bomb the convoy as the allies headed west. The payoff matrix for the options open to the opposing commanders is shown in Table 13.6.

From General Kenney's standpoint, if he sent his forces up the northern route and the Japanese chose to sail to the north side of the island, since weather was expected to be poor, his pilots would be able to bomb the Japanese for just two days. However, if after flying north in the foul weather the allied bombers discovered that the Japanese had taken the southern route, they would still be able to bomb enemy ships for two days. In contrast, if Kenney sent his forces south, only to find out that the Japanese were on the other side of the island, the allies would have to be content with only one day's bombardment. But if they were lucky and found that the Japanese had taken the calmer waters of the southern route, they would be able to bomb the convoy for three days. We can show the situation facing General Kenney in terms of a decision tree (see Figure 13.9).

FIGURE 13.9 Decision Tree for Allied Strategy in the Battle of the Bismarck Sea

You can see from the decision tree that General Kenney's best option was to follow the northern route. Adopting the minimax strategy (by which you minimize your maximum regret: one day of bombing for the allies and three for the Japanese), Kenney realized that the worst he could do by flying south was to bomb the Japanese convoy for one day. However, if he sent his planes north, the worst outcome for him was to inflict two days of damage. Therefore, his **dominant strategy** was to fly north. That is, that strategy provided him with the highest payoff given what he expected the Japanese to do. Similarly, for the Japanese fleet, its worst fears would have been realized if it had taken the calm waters of the south and had been quickly discovered by the allies for a full three days of attack. Hence, this classic battle situation presented the game theorists with a textbook case of a Nash equilibrium. In reality, the Japanese sent their ships along the northern route. General Kenney also went north, and his planes delivered a devastating blow to Japanese aspirations in the Pacific with two days of incessant bombing.

Game Theory in Local Government Decision Making

The most often cited examples of the use of game theory involve strategic decisions in military operations or international relations. This powerful analytical technique may also be used to aid local government decision makers. We have considered a hypothetical case of a city mayor confronting a labor union. In reality, local governments often face the prospect of strikes resulting from disputes with labor unions. Until about two decades ago, a universal ban prevented strikes by state and local government employees. For example, New York's Taylor law flatly prohibits strikes by public employees and prescribes the forfeiture of tenure and double loss of pay for striking workers.¹⁸ However, in recent years this stance has softened somewhat. By the early 1990s, about ten states permitted certain groups of government workers to strike under specific circumstances.

Needless to say, strikes or work slowdowns by public employees can cost cities millions of dollars. At the same time, a city can get into a financial bind by signing an overly generous pay package with the unions. In fact, many have blamed New York City's brush with bankruptcy in 1975 on politicians too eager to placate the demands of labor unions.¹⁹ Strikes usually do not help workers' unions, either. Unions lose out in terms of public support and may even risk court-imposed sanctions and fines.

If a union and a city administration are not careful, they can get locked into a losing proposition. Such a situation is known as a **prisoner's dilemma**. The name comes from a game in which, because of a lack of mutual trust, understanding, and cooperation, both players must settle for a less than satisfactory resolution. Let us consider the following scenario, portrayed frequently on television in police and courtroom dramas. You and your best friend have committed a crime, but no one saw you do it. The only way the district attorney can get a conviction is if either or both of you confess. Knowing this, the police have kept the two of you in different cells, with no possibility of communication. Your attorney tells you that you have two options: you can either confess or not confess. If you choose to confess, you will face two alternate scenarios. If your unrepentant, bullheaded friend does not confess, you will be rewarded with the witness protection plan while your friend languishes in jail for fifteen years. If, on the other hand, your weak and gutless friend decides to confess as well, both of you will draw five-year prison terms. If you decide not to confess and your friend betrays you, you end up with the fifteen-year prison term. However, if he turns out to be equally trustworthy and

TABLE 13.7 **The Prisoner's Dilemma**

Your strategy	Your friend's strategy	
	Confess	Do not confess
Confess	5, 5	Witness protection, 15
Do not confess	15, witness protection	0, 0

remains steadfast in not confessing, the prosecution has no case, and both of you are acquitted.

What should you do? Before answering this question, let us consider the payoffs for each strategy (see Table 13.7). You can see in the matrix that *collectively*, both of you will be better off choosing not to confess. However, you will reach this desirable position (from both of your viewpoints) only if you trust each other. Without trust, you will look at only two strategies: confess or do not confess. If you confess, the maximum risk is five years in prison. If you do not confess, the maximum risk is fifteen years in jail. Therefore, following the decision rule of minimizing your maximum loss (the minimax rule), you should choose to confess. Following the same rule, your friend will arrive at the same conclusion, leading to a joint confession and a five-year prison term for both of you. You may notice that in this case, confessing is the dominant strategy for both players.

In our example in Table 13.5, both the mayor and the union could have gone for a negotiated settlement in which both parties would have been better off. The labor union would have improved its position from 1 to 3 and the mayor from 1 to 2.

Game theory, by providing insight into the dynamics of a conflict, can help the involved parties avoid the prisoner's dilemma. The reason parties get locked in a prisoner's dilemma is that they do not trust each other. If the mayor and union had not been as mistrustful of each other, they could have found a mutually acceptable arbitration board. In many areas of decision making under uncertain conditions, an effort to develop trust can save everyone from being locked in an undesirable position.

The Golden Rules of Decision Making under Uncertainty

Identifying the optimal strategic moves in situations of uncertainty has been the subject of a great deal of academic discussion. On the basis of this lengthy discourse, I can offer four rules of advice:

Rule 1: Look ahead and reason back. Successful strategies depend on conscious forward thinking. However, while planning, one must think about the possible reaction of the opponent and the opposing course of action. The old Native American adage of learning to walk a mile in someone else's moccasins has considerable validity. Putting yourself in your opponent's shoes requires you to draw lessons from past experience to anticipate where your moves ultimately will take you.

Rule 2: If you have a dominant strategy, play it. It may not win you many friends, but if you have a dominant strategy, you must play it regardless of the situation. This rule is especially true in **zero-sum games**, in which the players' interests are strictly opposed and one player's gain is the other's loss. If we are playing poker, your win is

my loss, and the total amount to be distributed is fixed. The confrontation between the mayor and the union may be construed as a zero-sum game. Everyone must play the dominant strategy and, inevitably, must choose the undesirable outcome—deadlock. However, given the absence of trust, by playing the dominant strategy, each party ensures that it will not be stuck in an even worse position.

Rule 3: Successively eliminate all dominant strategies from consideration. If you have a number of possible alternatives, you should compare the outcomes and eliminate the dominant strategies. We followed this rule during our discussion of the decision tree. You may recall that by the process of elimination, we arrived at the most desirable strategy of building a levee and a flood wall. Proceeding in a similar way in your decision making may bring you to a unique alternative that is better than all others.

Rule 4: In the absence of a unique dominant strategy, look for an equilibrium. In the absence of a dominant strategy, settle for a point that represents the best response under the circumstances, given the players' positions. These circumstances can be quite complicated. Game theorists have pointed out that some situations may have more than one equilibrium position, or none at all. Since the discussion of these complicated scenarios is likely to take us beyond the scope of this book, for the sake of economy I will table these questions for now and simply refer the more curious readers (and certainly the adventuresome) to other books and articles.²⁰

STRATEGIES TO OVERCOME THE PRISONER'S DILEMMA

Uncertainty often causes us to take actions that are not in our best interest. This is true for individuals, organizations, and even nations engaged in international strategic moves. Often people get locked into situations that offer poor solutions; however, given the circumstances, there seems to be no other way out.

You can see that following the rules of rational decision making under uncertainty can put you in an undesirable situation. We will discuss some ways of building trust and engineering cooperation in the next section. However, often in life we must proceed without trusting our adversaries. In international relations, nations get into wasteful arms races (even those that can least afford to do so). By engaging in destructive labor disputes, industries around the world have been forced to shut down, causing unemployment for workers and economic hardship and loss of capital for owners and management. Of course, many interpersonal disputes remain unresolved because of lack of trust.

However, trust is not developed in a vacuum. It can evolve only through players' trustworthy actions. If I know an individual to be reliable, I can expect cooperative behavior in response to my cooperative gesture. However, this reciprocity of cooperative behavior can break down because of greed or fear. In our example of the prisoner's dilemma, I can truly win if my friend is made to believe that I will not confess—and then I turn star witness for the prosecution. If I pursue this strategy, I am guilty of greed. If I know that I can win with an unexpected defection, like all other swindlers and con artists, I will slowly develop trust and then defect at the most opportune time. Or I may defect out of fear. I may believe that you are about to break our bond of trust, so I may do it on my own. From violent ethnic strife in Eastern Europe and Africa to broken

relationships between friends and relatives, life is full of examples of such defections, which result in suboptimal payoffs for the players.

Greed and fear can be minimized if there is an avenue for punishing defection. The Mafia in southern Italy was able to maintain its iron grip over society by using *omertà*, the universal code of silence. Regardless of the situation, a defector was sure to face the fearsome wrath of the organization. This code kept the system of organized crime going for years. In the United States the court system punishes breach of contract, as defined by clearly understood and vigorously enforced tort laws. If the players are aware of the costs of defection, and the costs are higher than the expected gains, they can be assured of a cooperative game.

However, in many cases, such enforced cooperation is not viable. A great deal of effort has gone into discerning the best strategies under complete uncertainty, when cooperation between two players cannot be guaranteed. In the early 1980s Robert Axelrod, a University of Michigan professor of political science, organized a tournament.²¹ Game theorists from around the world submitted computer programs proposing strategies for dealing with uncertainty. These strategies were matched against one another in pairwise competition, in which each computer program was pitted against every other computer program. Each game was repeated 150 times. The strategies varied from nasty (defection on every move) to saintly (cooperation regardless of the opponent's action), from simple techniques to highly complex systems of calculated cooperation and punitive defection. The game that received the highest number of points was one that was submitted by Anatol Rapoport, a professor of mathematics at the University of Toronto. This surprising winner followed a simple strategy: tit for tat. It started out by cooperating (that is, not confessing, in the language of our example). After this it simply repeated the last move of its opponent. In other words, if the opponent cooperated, the tit-for-tat strategy did the same. However, if the opponent defected, the strategy sought immediate retribution and defection. But if the opponent, having defected, went back to cooperation, it was quick to forgive and adopt a cooperative position.

Axelrod attributed the success of the tit-for-tat strategy to four desirable qualities: clarity, "niceness," provocability, and forgiveness. The rules of this strategy are simple and, therefore, *clear*. The strategy is *nice* in the sense that it does not defect unprovoked. Although it is quick to exact retribution for a breach of trust (*provocability*), it is equally quick to *forgive* and go back to the cooperating mode. Axelrod claimed great possibilities for the tit-for-tat strategy and saw through it the evolution of a cooperative system. During the period of the renewed Cold War (after the speech in which President Ronald Reagan referred to the Soviet Union as an "evil empire"), this simple strategy symbolized for many the hope of a new, trusting world order.

However, despite the early hopes for greatness, later scrutiny of the tit-for-tat strategy found damaging flaws in its universality. First, the strategy was the winner only in terms of cumulative score; it could not beat any strategy in a pairwise game. The best it could do was to tie with the cooperative (saintly) games, or the games following strategies close to its own, starting with cooperation. It won overall because, since it echoed the opponent, it always came close, regardless of strategy, a quality that could not be matched by any other strategy.

The tit-for-tat strategy also suffers from some other serious flaws. Avinash Dixit and Barry Nalebuff point out that if two opponents are playing this strategy and one defects

by mistake, or the move is seen as a defection, the tit-for-tat strategy will kick in, and the two will not be able to extricate themselves from a never-ending cycle of retribution. Instead of tit for tat, Dixit and Nalebuff suggest a more forgiving strategy.

OTHER STRATEGIES: TRUST AND BARGAINING

It is obvious that no one would want to get mired in a game of “getting even.” Mahatma Gandhi often quipped that the strategy of “a tooth for a tooth and an eye for an eye” ultimately leaves contenders toothless and blind. Across the United States a great deal of academic energy is being devoted to finding ways to resolve seemingly intractable disputes. The key to the resolution of conflict lies in developing trust. When trust is lacking, participants need external enforcement of compliance through the imposition of costs for noncompliance, and as always in a conflict situation, they should rely on bargaining. In their widely read book *Getting to Yes*, Roger Fisher and William Ury of the Harvard Negotiation Project spell out a number of strategies to achieve agreement in disputes.²² Although these strategies make for extremely useful and lively debates—and I strongly recommend that you become familiar with them—their full discussion falls outside the scope of this book.

I leave you with a final word of caution. Because of the very nature of disputes, not all of them have peaceful resolutions. For many different reasons, the contenders may deem open hostility to be more desirable than a negotiated compromise. In such cases, the only outcome is continuing conflict.

KEY WORDS

Actions (p. 323)	Outcomes (p. 323)
Collectively exhaustive (p. 324)	Prisoner's dilemma (p. 337)
Decision tree (p. 322)	Risk (p. 320)
Dominant strategy (p. 337)	Risk averse (p. 331)
Expected payoff (p. 321)	Risk lover (p. 331)
Game theory (p. 332)	Risk neutral (p. 331)
Minimax strategy (p. 334)	Uncertainty (p. 320)
Mutually exclusive (p. 324)	Zero-sum games (p. 338)
Nash equilibrium (p. 335)	

EXERCISES

1. What is an expected payoff? How does the concept of expected payoff help analyze a situation of uncertainty? In this context, explain “dominant strategy” and the minimax rule.
2. What is a zero-sum game? With an appropriate example, explain how conflict arises in the area of resource allocation (resulting from a public policy) because of a zero-sum situation.

3. What is the prisoner's dilemma? What are its outcomes for the participants? How can it be prevented in real life?
4. In July the parks and recreation department in your town wants to organize an exhibition of local arts and crafts. The department is considering whether to hold the exhibit inside the downtown sports arena or outside in its open-air parking lot. However, the problem is that while planning the location of the exhibition, the planners face uncertainty about the weather. The weather bureau reports that there is only a 20 percent chance of rain. If the town holds the event inside the sports arena, after paying for the use of the facility, it expects to break even if it rains. If it does not rain, the town will make a profit of \$15,000 from the event. On the other hand, if the exhibition is held outdoors, it can be either a great success or a real failure, depending on the weather. If it does not rain, the town will make a tidy profit of \$35,000, but if it rains, it stands to lose \$10,000. Draw a decision tree and explain the options open to the decision makers. What would be your recommendation?
5. Your town is located on the shores of the Atlantic Ocean, directly in the line of devastating hurricanes. You are evaluating a policy to invest \$10 million in emergency preparedness. It is estimated that the chance that a moderate to strong hurricane (categories 2 and 3) will come your way next year is 30 percent. The chance of being hit by a devastating category 4 hurricane is 5 percent. Right now, without this additional preparation, the town can handle small tropical storms without much problem or property loss. If there is a category 2 or 3 hurricane, however, the estimated loss will be about \$15 million, and for a fierce category 4 storm, the estimated property loss could rise as high as \$50 million.

The financial administration department estimates that with the \$10 million invested, the town will be able to withstand a strong hurricane of category 2 or 3 with a minimum damage of \$5 million. Even if there is a giant hurricane of category 4, the damage estimate goes no higher than \$20 million. Should the town invest the money in disaster preparedness? Draw the decision tree and write a report explaining your recommendation.

6. Suppose you are an aide to the governor of your state. Because of a much publicized story of a mentally deranged person killing a number of innocent victims in a crowded restaurant with an automatic weapon, the governor is considering a proposal to ban all sales of automatic guns within state borders. This issue is emotionally charged, and the governor is keenly aware of the political cost of an unpopular decision.

Facing this controversial problem, the governor sees three options: to do nothing, to propose a mild law banning the sale of a few such weapons, or to establish a commission. If the governor does nothing, there would be a political cost, to which the governor's advisers assign a value of -3 . If a mild ban is unilaterally imposed, it will cause a net loss of popularity, valued at -2 . On the other hand, if an independent commission is set up, the governor may be able to circumvent this lose-lose situation. However, the governor has no control over the recommendations of an independent commission. The commission may come up with a "do nothing" recommendation, which would absolve the governor for not doing anything (value = 0). Another possibility is that the commission will recommend a mild ban, or it could suggest a strict ban on the ownership of automatic weapons. The governor, however, retains the right to go along with the commission or to reject

its recommendations. If the commission recommends a mild ban, the governor can accept it for a political payoff of +3, or he may reject it and do nothing for a small loss of -1 (the commission has taken the edge off the governor's inaction).

A real problem may arise if the commission recommends a radical plan of gun control. If the governor accepts it, he will be hit for a loss of -10, but if he rejects the plan, he will incur a loss of -4. The probability that the commission, which will have broad-based support, will recommend doing nothing is .3, the chance that it will suggest a mild ban is .5, and the chance that it will support a strict gun control measure is .2. If you are a political analyst for the governor, what will be your recommendation and why?

7. Many social conflicts arise when rival factions place themselves in a prisoner's dilemma situation. With an appropriate example, discuss the prisoner's dilemma, pointing out the reasons for the intractability of problems and suggesting some possible measures that may help generate cooperation among the parties involved.

Notes

1. Frank Knight, *Risk, Uncertainty, and Profit* (Boston: Houghton Mifflin, 1921).
2. Over the years, authors have altered the definitions of risk and uncertainty proposed by Knight. You may come across a definition of risk that describes it as associated with a project itself, whereas uncertainty is thought to relate to the overall environment. However, in a delightful work on the subject, Young Back Choi argues for the original definitions of these terms. Following Choi, I stick to Knight's definitions of risk and uncertainty. See Choi's *Paradigms and Conventions: Uncertainty, Decision Making, and Entrepreneurship* (Ann Arbor: University of Michigan Press, 1993).
3. Choi, *Paradigms and Conventions*, 13–16.
4. Dipak K. Gupta, "Economics and Collective Identity: Explaining Collective Action," in *The Expansion of Economics and Other Disciplines: Toward an Inclusive Social Science*, ed. Shoshana Grossbard-Shechtman (Armonk, N.Y.: M.E. Sharpe, 2001). A number of eminent economists have broken ranks with the traditional view of utility maximization. See George A. Akerlof and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics* 115 (2000): 715–753; Timur Kuran, "Ethnic Norms and Their Transformation through Reputation Cascades," *Journal of Legal Studies* 27, pt. 2 (1998): 623–659; Amartya K. Sen, "Goals, Commitment, and Identity," *Journal of Law, Economics, and Organization* 1 (1985): 341–355; and Howard Margolis, *Selfishness, Altruism, and Rationality: A Theory of Social Choice* (Cambridge, United Kingdom: Cambridge University Press, 1982).
5. Harvey Leibenstein, "The Prisoner's Dilemma in the Invisible Hand: An Analysis of Intrafirm Productivity," *American Economic Review* 72 (1982): 92–97.
6. Choi, *Paradigms and Conventions*, 15–16.
7. Arthur M. Silverstein, *Pure Politics and Impure Science: The Swine Flu Affair* (Baltimore, Md.: Johns Hopkins University Press, 1981). Also see Richard E. Neustadt and Harvey V. Fineberg, *The Swine Flu Affair: Decision-making on a Slippery Disease* (Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1978).
8. For an actual study of the probability of flood damage, see *Flood Proofing: How to Evaluate Your Options: Decision Tree* (Fort Belvoir, Va.: U.S. Army Corps of Engineers, National Flood Proofing Committee, 1995).
9. For a comprehensive review of the problem, see Mark Baldassare, *When Government Fails: The Orange County Bankruptcy* (Berkeley and San Francisco, Calif.: University of California Press and Public Policy Institute of California, 1998). In this section, I concentrate mostly on

- risk management. However, a number of studies, in addition to identifying the problem of assuming undue risks, point to fraud and criminal activities in this expensive fiscal fiasco. See, for example, Susan Will, Henry N. Pontell, and Richard Cheung, "Risky Business Revisited: White-Collar Crime and the Orange County Bankruptcy," *Crime & Delinquency* 44 (1998): 367-387.
10. For an excellent discussion of derivatives, see Julian Walmsley, *New Financial Instruments*, 2nd ed. (New York: Wiley, 1998).
 11. See Associated Press, "Former Orange County Treasurer Had Brain Damage, Witnesses Say," *San Diego Union-Tribune*, November 19, 1996.
 12. For an informative, highly entertaining explanation of game theory, see Avinash K. Dixit and Barry J. Nalebuff, *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life* (New York: W. W. Norton, 1991).
 13. David M. Kreps, *Game Theory and Economic Modelling* (Oxford, United Kingdom: Oxford University Press, 1990). Also see Jeremy I. Bulow, John D. Geanakoplos, and Paul D. Klemperer, "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy* 93 (1985): 488-511.
 14. Roger Fisher and William Ury, *Getting to Yes: Negotiating Agreement without Giving In* (New York: Penguin, 1981). Also see Howard Raiffa, *The Art and Science of Negotiation* (Cambridge, Mass.: Belknap, 1982).
 15. *Newsweek* magazine, in reviewing Fisher and Ury's *Getting to Yes* (see note 14), called the book "a coherent 'win-win' negotiation which, if it takes hold, may help convert the Age of Me to the Era of We" (quoted on the book's back cover).
 16. Bruce Bueno de Mesquita, *Principles of International Politics: People's Power, Preferences, and Perceptions* (Washington, D.C.: CQ Press, 2000), 48.
 17. For an excellent discussion of the importance of game theory, see John L. Casti, *Five Golden Rules: Great Theories of 20th-century Mathematics—and Why They Matter* (New York: Wiley, 1996).
 18. Robert J. Thornton, "Unions and Collective Bargaining," in *Management Policies in Local Government Finance*, ed. J. Richard Aronson and Eli Schwartz (Washington, D.C.: International City/County Management Association, 1996), 418.
 19. "Message to Rudy," *The New Yorker*, May 15, 1995.
 20. For a discussion of multiple equilibrium, see John C. Harsanyi, "Advances in Understanding Rational Behavior," in *Foundational Problems in the Special Sciences*, ed. Robert E. Butts and Jaakko Hintikka (Dordrecht, the Netherlands: D. Reidel, 1977). For an example of this kind of situation and its resolution, see Frank C. Zagare, "Rationality and Deterrence," *World Politics* 42 (January 1990): 238-260. Again, for the most readable explanation, see Dixit and Nalebuff, *Thinking Strategically*.
 21. For the results of this tournament, see Robert Axelrod, *The Evolution of Cooperation* (New York: Basic Books, 1984).
 22. Fisher and Ury, *Getting to Yes*.

14

Choosing the Best Alternative: Cost-Benefit Analysis

Your city is considering acquiring a parcel of vacant land near the center of the city to develop a new convention center with a hotel complex. Your job, as a policy analyst, is to evaluate the proposal and send your recommendation to the city council. The question is, of course, how to go about it.

It is obvious that you will recommend the project if the benefits outweigh the costs. When we make a decision, any decision, we consciously or unconsciously evaluate its potential benefits and costs. In fact, this evaluation process is so fundamental to human cognition that many economists have equated it with the very notion of human rationality.¹ Economists argue that when individuals make decisions, even emotional ones, they have a good sense of the benefits and costs of their actions. How do we choose our mates in marriage? Poets and authors of romantic novels claim that love is blind—anyone can fall in love with anyone else without following any definite pattern of behavior. Yet statistical data show that marriage partners tend to match each other's "endowments" (age, looks, wealth, education, social standing, and so forth). That is, their actual preferences reveal a "rational choice" based loosely on the economic notion of **cost-benefit analysis**.² Thus, you will probably not be surprised that this chapter on cost-benefit analysis mirrors chapter 5 on critical thinking. Cost-benefit analysis, which emphasizes monetary evaluation, provides critical thinking with a more formal structure. I mentioned that the techniques of statistics and operations research greatly aid the process of thinking critically. Hence, because cost-benefit analysis draws on all these tools, I am discussing it at the very end of our discussion of analytical techniques.

In sum, the fundamental principles of cost-benefit analysis are the following:

- *When considering a single project, accept it if its benefits are greater than its costs.*
- *When considering alternative projects, choose the one that gives you the highest benefits in relation to the costs.*

In our convention center example, we should go ahead with the project if its benefits are greater than its costs. Similarly, we should evaluate alternative uses for the land (for example, a new public library, a public park, or a shopping mall) with regard to their respective benefits and costs and choose the one that gives us the highest benefit relative to the costs.

If you recommend building a new convention center, how much should the city pay for it? Those who derive their livelihoods from the tourist industry may benefit more than those who do not. Thus, those in the former group may be willing to pay more for it than others. If the total amount the city is willing to pay (reflecting the total utility of the project to its stakeholders) is greater than the cost of the project, it is worth undertaking. This total amount is known in economic terms as the **consumer surplus** (see chapter 3). Therefore, this question goes to the heart of cost-benefit analysis. However, the more we look into the issue, the more complicated it becomes. To fully appreciate the complexity of the problem, we start with the process of conducting a cost-benefit analysis.

Similar to the familiar process of critical thinking, the process of conducting a cost-benefit analysis is as follows:

1. Define the goal(s) of the project.
2. Identify the alternatives.
3. Make an exhaustive list of all benefits and costs, present and future.
4. Estimate and express benefits and costs in monetary terms.
5. Forecast the future streams of benefits and costs, if needed.
6. Choose the alternative with the largest benefit in comparison with the cost.

Let us now discuss in detail each of these six steps. This powerful analytical tool can be used to evaluate almost every social decision. And as you will see from our discussion, its thoroughness (which is directly linked to the cost of the study) must be matched against the importance of the project. If we want to evaluate a proposal for a neighborhood park, we may want to devote a modest amount of time and effort. In contrast, the evaluation for the construction of an atomic power plant, which may affect the health and welfare of hundreds of thousands of residents in the region, would require a much greater degree of precision.

SOCIAL VERSUS PRIVATE COST-BENEFIT ANALYSIS

Before getting to the steps mentioned above, we should briefly discuss the various ramifications for the public versus the private sector. An example may clarify the distinction. An analyst friend, who works for a small neighboring town, called me to say

that he was disappointed. The town was negotiating with a national retail chain store to locate within its boundaries. After months of negotiation, the store decided to locate in a more affluent medium-sized city next door. Because my friend's town was coming out of a long fiscal slump and the store would have given it a significant boost, he complained, "Didn't the management of the retail chain know how much social good they could have achieved by choosing to locate here?"

Indeed, if you considered the positive externalities of this relocation decision, you might have agreed with him. However, from the standpoint of the retail chain, the private net benefit was the only calculation that counted. This difference in perception explains the distinction between private and public cost-benefit analysis. Every project, large or small, contains social benefits and costs. Retail stores generate employment, produce sales and income tax, and may even increase property values in the surrounding areas. In contrast, they may also create traffic congestion and pollution. When conducting a cost-benefit analysis for a public project, we must take into account these **positive externalities** and **negative externalities**; because private firms are not rewarded for the positive externalities they generate in the community, nor penalized for negative externalities (unless their actions cause a lawsuit), they do not consider them in their cost-benefit calculations.

DEFINING GOALS

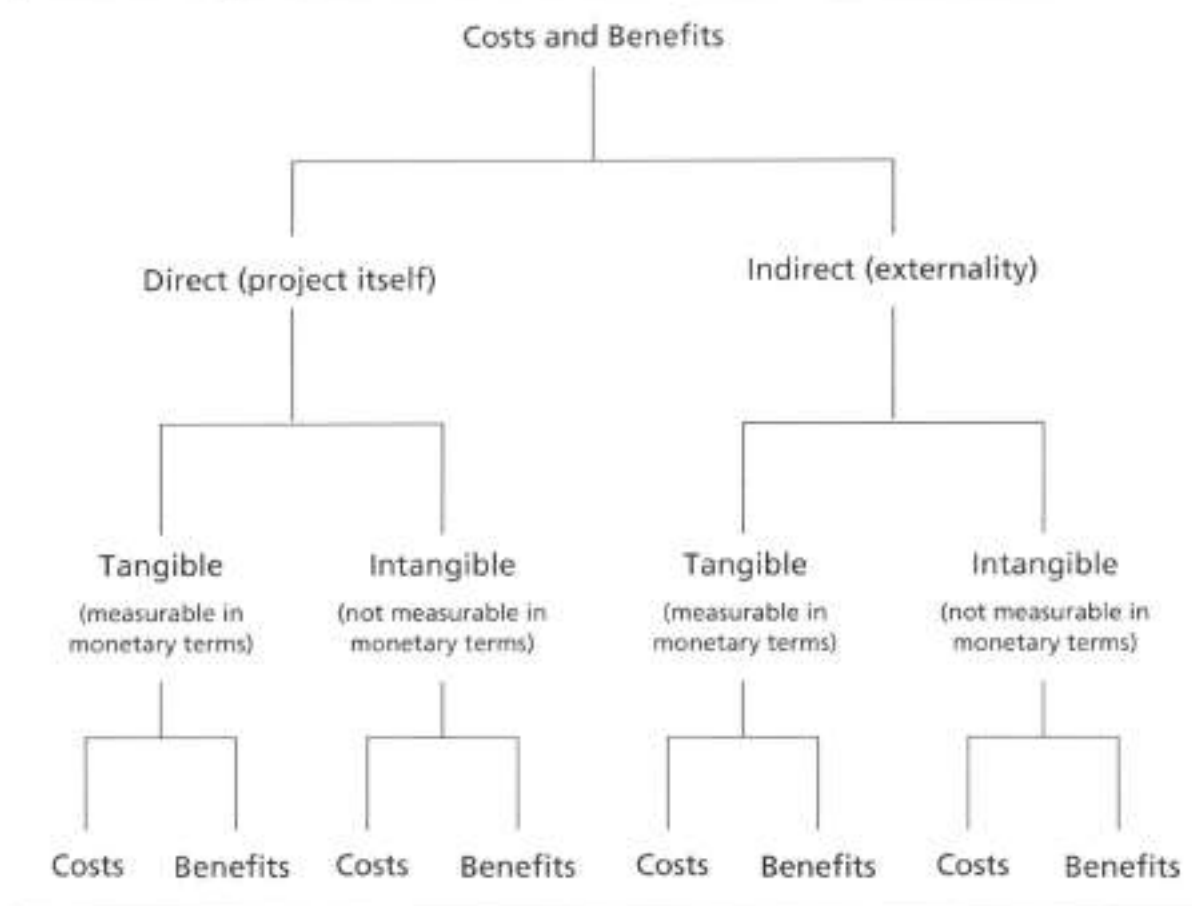
The first task of a cost-benefit analysis is to identify the goals of the project. The clearer the goals are, the easier it will be for an analyst to select the best course of action for achieving them. In chapter 5, I demonstrated the importance of defining the goals of projects. When evaluating the convention center proposal, we must determine whose goals to maximize, the city's, the county's, or the convention center authorities'. We also need to know the goals of the project. We may identify as a goal an increase in tax revenue, urban renewal (as a result of the new convention center), or the enhanced image of the city. Elected representatives may set the goals, or we may determine the wishes of the stakeholders through surveys and focus groups.

IDENTIFYING ALTERNATIVES

The second step in preparing a cost-benefit analysis is to identify the alternatives. Typically, when there are not many identifiable alternatives, we evaluate a project against the option of doing nothing. However, if there are other feasible uses of the site, such as a new public library, a park, or a shopping mall, we should consider them in our analysis.

LISTING THE COSTS AND BENEFITS OF THE ALTERNATIVES

After selecting the alternatives, one should make as exhaustive a list as possible of the costs and benefits of the various alternatives. It may be useful to use the scheme shown in Figure 14.1 to delineate costs and benefits. They can be broadly classified into two

FIGURE 14.1 Classification of Costs and Benefits

categories: direct and indirect. The **direct costs and benefits** are those that are associated directly with the project itself. The **indirect costs and benefits** are those that affect the surrounding community but do not show up in the ledger of the project.

In our convention center example, the revenues from the new facilities are considered to be the direct benefits, and the construction costs are the direct costs of the project. However, beyond these benefits and costs are the externalities of the proposed project. Thus, the generation of economic activities, such as increased business for the surrounding areas, is to be counted as an indirect benefit. The increased activities, however, may create factors that are detrimental to the city, such as increased traffic, pollution, or crime. These are the social or indirect costs of the project.

Not all costs and benefits are measurable in monetary terms. Many costs and benefits are primarily qualitative in nature and as such cannot be readily expressed in dollars and cents. For example, a beautiful convention center can produce the intangible benefit of newfound pride in the city. Its construction can bring about a change in attitude in the city's citizens. In contrast, such negative conditions as environmental hazards brought about by the construction of the hotel and convention center would constitute part of its social costs. These are the **intangible costs and benefits** of a project. Analysts should be as comprehensive as possible in enumerating the costs and benefits of a public project, both intangible and tangible. I present a list of possible tangible and intangible costs and benefits in Table 14.1.

TABLE 14.1

Benefits and Costs of Convention Center

	Benefits				Costs	
	Direct		Indirect		Direct	
	Tangible	Intangible	Tangible	Intangible	Tangible	Intangible
Revenues from the convention center	Appearance of the new facility		Increased business for the surrounding areas	Increased civic and community pride	Construction expenses and other related costs	Increased costs to business as a result of traffic congestion
			Higher property, sales, and transient occupancy taxes for the city		Cost of acquiring land	Increased crime and pollution
			Increased property values for some neighborhoods			Environmental degradation
						Displacement of the poor from the neighborhood

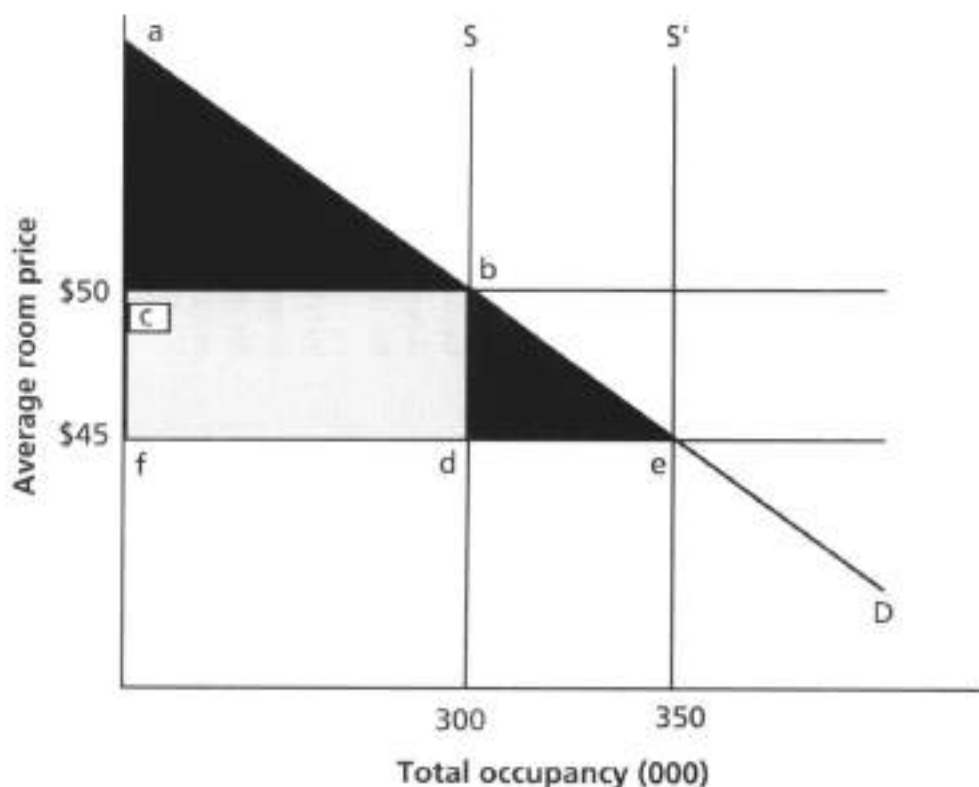
ESTIMATION AND VALUATION OF BENEFITS AND COSTS

One of the most difficult problems of conducting a cost-benefit analysis is that many of the benefits and costs may not be measurable in monetary terms. Another problem is that if they are to be accrued in the future, they must be estimated. It is not enough to state that the construction of a new convention center would increase revenue for the city. We need to come up with a reasonable estimate of how much additional revenue there would be.

A convention center would draw groups of people from outside the region for business meetings. The proposed center would provide large rooms for various gatherings as well as hotel space for convention participants and their guests. Suppose that city hotels can accommodate 300,000 guests per year. The new convention center would increase this capacity to 350,000. The average cost of renting a room in the city is \$50, but the increased supply of rooms would reduce this cost to \$45. (see Figure 14.2).

You may recall our discussion of consumer surplus from chapter 3. We illustrate this concept in Figure 14.2, in which the demand curve for hotel rooms (D) is shown as a downward-sloping heavy line. Suppose that before construction of the convention center, the supply curve was vertical line S , which gave us an equilibrium occupancy rate of 300,000 rooms at an average room cost per room of \$50. At that point, the triangle abc represented the total benefit, or consumer surplus. As a result of the convention center, the supply increased to S' , with an equilibrium room occupancy rate of 350,000 and an average rental rate of \$45. As you can see, this increase in supply and reduction in price

FIGURE 14.2 Consumer Surplus and Social Benefits



have allowed 50,000 extra guests to visit the city. This new situation has enlarged the area of total consumer surplus to the triangle *aef*.

Of the total consumer surplus represented by triangle *aef*, the part described by triangle *abc* is not new. However, the convention center project has in fact created two new areas of consumer surplus, the rectangle *cbdf* and the triangle *bde*. Rectangle *cbdf* is an added surplus to consumers. That is, as a result of the reduced price of lodging in the city, the gain to consumers is a dollar-for-dollar loss to producers—the hotel and motel owners who had to lower their rates to fill the additional capacity. Rectangle *cbdf* illustrates what microeconomists call the **pecuniary effect**, which occurs when a change in the welfare of one group of individuals comes at the expense of some other group. Since the gains of the gainers exactly match the losses of the losers, for society as a whole, there is no change in welfare, unless we want to value one group's gain differently from the other group's loss. In that case, society must make a value judgment about the redistribution of income. We will discuss this issue later in the chapter.

Returning to Figure 14.2, we see that the true additional consumer surplus is represented by triangle *bde*, the area of **net social benefit**. Using the Pythagorean theorem, we can calculate the area of triangle *bde* as

$$\frac{\text{Height} \times \text{Width}}{2} = \frac{(\$50 - \$45) \times 50,000}{2} = \$125,000$$

Thus, the net gain in consumer surplus to the city is \$125,000. Let us suppose that the project has been financed with municipal bonds. The bonds cost city taxpayers \$75,000 per year. Therefore, the net gain to the city is \$50,000, the difference between the gain in consumer surplus and the cost of servicing the loan (\$125,000 – \$75,000 = \$50,000).

The preceding problem was an easy one to solve. In real life, calculations of costs and benefits are rarely as simple as those in our example. Table 14.1 gives you an idea of the different kinds of costs and benefits that have to be estimated and translated into monetary terms.

One of the most important aspects of conducting a cost-benefit analysis is the valuation of intangibles, which are not bought and sold in the market. Yet for the sake of comparison, an analyst must ascribe monetary values to these items. Let us consider a few examples.

Can We Put a Price Tag on the Intangibles of Life?

How would you value the life of a human being, or the risk of injury that could result in physical disfigurement, or irreparable damage to the environment? Putting a value on such matters evokes controversy, yet it has become a routine matter. For example, when a jury hands down an award for pain and suffering, loss of face, disability, or loss of life, it is imputing a value to the most intangible aspects of life. The case of *State of Alaska v. Exxon Corporation*, resulting from an oil spill in the pristine and ecologically fragile Prince William Sound, is an example of putting a specific monetary value on the loss of habitat. It is indeed legitimate and necessary to consider ways of imputing monetary value to these intangibles when conducting a cost-benefit analysis.

For instance, our large construction project carries the possibility of accidents resulting in severe injuries and even death. In such circumstances, it is essential to include the costs of such accidents. There are several methods for calculating these costs: the **face value of life insurance**, **discounted future earnings**, and **required compensation**. The face value of life insurance measures the monetary worth of one's life by the amount of one's life insurance policy. However, people buy insurance for many different reasons (for example, to some, it is a form of forced saving), and as such, their purchase may have little to do with their perception of the value of their own lives. In the discounted future earnings approach, often used in court cases, a person's life is worth the discounted value of future income. Future earnings are discounted because a dollar in the future is worth less in today's money. But since it evaluates life by one's earning potential, this approach undervalues the lives of those individuals whose talents are not sold in the market or who have stopped earning money. Therefore, it will fail to place much value on homemakers, retirees, and people with disabilities who cannot work.

The market valuation of life does not address the crucial question of the worker's perception of the extent to which the added risk is compensated by the income differential. The answer would imply how much an individual believes his or her life is worth. There are people whose jobs carry almost no risk of death (for example, elementary school teachers, bank clerks), whereas the jobs of others carry an inordinate amount of risk (for example, firefighters, members of a bomb squad). Therefore, if we take an individual's education, age, experience, and other relevant factors of earning determination as constant, we will arrive at a larger payment for the risky jobs to compensate those workers for their added risk. The calculation of this margin of compensation for risk is called the required compensation principle. This is an important issue, and many economists have attempted to estimate the size of this margin. From their studies it seems that this value varies from a lower boundary of \$2.5 million to \$5 million in 1988 constant dollars. This value turns out to be, on average, five to ten times the value of life calculated under the discounted future earnings principle.³

As you can imagine, the imputing of a very high value for human life would make many projects less than economically viable. Therefore, you may think of these numbers as quite excessive, until you realize that an individual who took such a job might not have been totally aware of the risk involved. This is particularly true in the field of high technology; for example, workers whose job was sealing radiation chambers in the construction of atomic power plants complained that they were not adequately apprised of the risk by management. Even individuals who are aware of risk may not have the bargaining power to gain adequate monetary protection against the loss of life. Finally, this measure does not take into account the externalities of such a loss. The death of an individual can destroy a family and cause irreparable damage to the welfare of those who were dependent on this person for financial and emotional security. In light of these kinds of externalities, the U.S. military often exempted only sons from the draft or ensured that two brothers did not serve on the same ship.

As technology improves we come to realize the deleterious effects of substances such as asbestos, whose risks most people were not aware of until many years later. In December 2000 the Environmental Protection Agency (EPA) announced its decision to clean up a dangerous chemical, PCB, from the waters of the upper Hudson River by dredging 2.65 million cubic yards of sediment along a forty-mile stretch. General

Electric had discharged an estimated 1.1 million pounds of PCBs into the river before 1977 from capacitor plants in Fort Edward and Hudson Falls, about forty miles north of Albany. As a result, the EPA claimed that a 200-mile stretch of the river, down to New York City, was contaminated. The project would cost GE an estimated \$460 million. However, the effects of PCBs had not been previously known, and therefore they were not banned before 1977.⁴

An example may help you understand the differences among the three methods of valuing a life. Suppose an innocent thirty-five-year-old schoolteacher is killed by the police during a high-speed car chase. Pursuing a suspect, a police cruiser goes through a red light and crashes into the teacher's car. The city is asked to pay compensation for this loss of life. Assume that the young man was earning \$35,000 a year and had purchased a life insurance policy worth \$150,000. According to the face value of life insurance method, the city will be liable for \$150,000. However, if we assume that the young man would have lived for another thirty years and earned his current salary, his lifetime earnings discounted at a 6 percent rate turn out to be \$201,022.19 (see the discussion of present value later in this chapter). In contrast, since he was not in a hazardous job, the required compensation method would estimate the value of his life at several million dollars.

Since valuation of most of the intangibles in life is often highly subjective, the numbers can vary, climbing to absurd amounts. In 1991, two years after the Exxon tanker *Valdez* had spilled oil, causing extensive environmental damage in Alaska, Exxon agreed to settle criminal and civil complaints brought by Alaska and the federal government for \$1.25 billion. Yet within a relatively short period, a study commissioned by the state and federal governments put the damage to the ecology at \$15 billion.⁵

Indeed, as a society we may at times place extremely high prices on projects. If a project threatens a species with extinction or destroys a place of national interest or veneration, then we may assume that the cost of its destruction is too high for any conceivable monetary compensation. To prevent the extinction of spotted owls, the U.S. government declared a moratorium on logging in Oregon in 1991.

How Can We Measure Future Loss or Gain?

The prospect of future loss poses one of the most difficult obstacles to public projects. In popular terminology, this is the dreaded NIMBY (not in my backyard) factor, which community groups can effectively use to stop the construction of projects that have widespread indirect benefits but impose specific costs on a certain community. Thus, while the construction of a new airport may prove to be a boon to a region's economy, the question remains as to which community will have to live with the noise and increased traffic. Although small in proportion to the total gain to the region, the cost of increased noise can have disastrous effects on property values in nearby neighborhoods.

An analyst is often faced with estimating a loss of property value that has not yet occurred. This estimate can typically be carried out using a causal regression model, discussed in Appendix B. We can form a regression model in which the price of property will be a function of changes (Δ) in

$$\text{Price} = f \Delta (\text{Noise, Pollution, Travel time, Other factors}) \quad (14.1)$$

In this case, we hypothesize that the price of property will depend on the altered levels of noise (measured in decibels) and pollution (measured by various

standardized emission units), which will have a negative effect on the price. A decrease in travel time (measured in terms of minutes to the airport), in contrast, is likely to increase the price. Other factors include the property's size, location, view, and so forth. For the model we can take these factors as given because the construction of the airport will not change them.

Taking a cross-section of city properties, we can estimate the relevant coefficients for noise, pollution, and travel time. The coefficients for each term measure the impact on the dependent variable of a one-unit change in the independent variable. By multiplying the estimated coefficients with the expected change in that variable as a result of airport construction, we can estimate the total loss to the property. Suppose our regression coefficient for the effect of noise on the price of property turns out to be $-\$5,000$. This would mean that a one-unit increase in decibel level would reduce the price of a piece of property by $\$5,000$. Suppose the environmental impact statement estimates that the new airport would add five decibels to the already existing noise level of a particular neighborhood. We can estimate the loss of property value for that neighborhood to be $-\$5,000 \times 5 = -\$25,000$. Other coefficients can be used in a similar manner to measure the total impact the new airport would have on property values.

This kind of estimation poses many problems. Property owners are likely to dispute the results, because the estimates certainly will not cover all the costs associated with increased noise and other kinds of pollution (such as the effect on the physical and psychological health of the residents). It is interesting to note that frequently there exists an asymmetry in information between the gainers and losers of large public projects. In some instances a small group of potential losers tends to know and care about its losses a lot more than the larger group of potential gainers. In such cases, well-organized groups are often able to stop a project through political protests or obstructive legal actions. Conversely, in other cases, in which the potential for individual gains is strong, a handful of powerful interest groups are able to get approval for a project that may inflict costs on a wide segment of society.

It should be obvious by now that inferring the value of nonmarketable items is not an easy task and often creates controversy. Yet as an analyst you may have to estimate the value of time saved as a result of a traffic diversion or the emotional cost of destroying a community to build a freeway through it. You have to approach such matters boldly but with caution. For example, a recent report suggested that the construction of high-occupancy vehicle lanes (the highway lanes set aside for vehicles carrying more than a certain number of passengers) on the perennially clogged Atlanta freeways reduced commuting time by fifteen minutes. You may be tempted to put a value on the time saved by multiplying it by the average wages of the commuters multiplied by their number, until you realize that the time saved is not likely to increase the commuters' working hours. Instead the fifteen minutes that would have been spent sitting in a traffic jam will now be spent pursuing enjoyable activities that carry no commercial value. You may instead consider the amount of gasoline saved by having to run the car engine for fifteen minutes and then calculate the money saved by commuters. In addition, you may look for the environmental benefits of reduced auto exhaust emissions.

In the previous pages we discussed the problems of putting monetary values on nonmarketable items. If you find items that are simply not translatable in money, you may do well not to overstretch your imagination. As we have seen, unless you are careful,

the valuing of nonmarketable items can quickly veer toward the ridiculous. Therefore, in such cases, an analyst should report accurately the intangible effects of the proposed project so that political decision makers can make informed decisions.

INTRODUCTION OF TIME: PRESENT VALUE ANALYSIS

In the preceding example of the convention center, we have a relatively simple choice to make on the basis of a onetime, lump-sum net benefit. However, the benefits and costs of most projects do not occur at one time. Instead, they come in over a period of time. This inclusion of time adds one more dimension to our problem.

To begin with, a dollar received a number of years down the road may not be worth as much as a dollar already in our pockets. I am always reminded of a local television commercial for an annuity program. The announcer asks viewers to join a "millionaires' club." If a young adult saves a certain amount of money per month, at the end of nearly thirty-five years, this individual will receive \$1 million from the annuity. Of course, during the dreamy announcement part of this commercial, the camera lens pans over all the trappings that are commonly associated with the lives of millionaires—a fancy home, a limousine parked in the driveway, and so on. Ask yourself, though, would \$1 million thirty-five years from now be worth \$1 million in today's money? Obviously, the answer is that the two amounts of money are not equal. But the question of the difference between a dollar in my pocket today and one in the future can be answered only if we understand the process of discounting.

To explain the process of discounting, I must first explain the process of compounding. Suppose I have invested \$100 in a certificate of deposit, maturing at the end of the year, at a 10 percent interest rate. At the end of the year, I will receive \$110. Thus, \$100 invested at 10 percent interest for a year will yield $\$100 \times (1 + 0.1)$, or \$110.

The preceding formulation is perfectly obvious. If I keep this investment one more year at the compounding interest rate of 10 percent, at the end of the second year, I will get back not another \$10, but \$11, because I will earn interest on the previous year's interest. Therefore, at the end of the second year, I will receive $\$110 \times (1 + 0.1)$, or \$121. By inserting into the preceding equation the formula by which we obtained the result of \$110, we get $\$100 \times (1 + 0.1) \times (1 + 0.1)$, or \$121, which can be rewritten as $\$100 \times (1 + 0.1)^2 = \121 .

If you are observant, you will note that keeping the money for two years requires us to multiply the original amount of money invested by 1 plus the interest rate (10 percent, or 0.1 in this case), the quantity raised to the power of 2. Therefore, if I had kept the money for three years, the exponent of the term within the parentheses would have to be raised to 3. By extending this logic, we can generalize by stating that the original investment of P_0 at a rate of interest of r percent for n years will give us P_n amount of money:

$$P_n = P_0 \times (1 + r)^n \quad (14.2)$$

where P_n is the principal amount at the end of the n th period, P_0 is the original principal amount at period 0, and r is the rate of interest.

Using a calculator, we can determine that \$155, invested at a 6.5 percent rate for seventeen years, would yield $\$155 \times (1.065)^{17}$, or \$452.14. This is the formula for the

computation of compound interest; it tells you how much a dollar invested today at a certain interest rate would be worth in the future.

In contrast, a dollar in the future may not be worth its full face value in today's currency; the forces of inflation, uncertainty, risk, and the plain fact that you would rather have your money now than at a later date may eat away much of its value. In other words, I may pose the question from the opposite direction: how much would a dollar be worth to you in the n th year in the future? In such a case, without compounding your initial investment, you would have to discount your future income. Let us take a specific example. Suppose I were to receive \$100 a year from today. Since I will be getting the money in the future, if I use a **discount rate** (which measures the intensity with which I want my money in the present) of 10 percent, the \$100 will be equal to

$$\frac{100}{1.1} = \$90.91$$

As in our previous example, if we are considering n years in the future, our future gain will have to be discounted by 1 plus the rate of discount, raised to the number of years we have to wait for the money. Thus, we can generalize the formula as

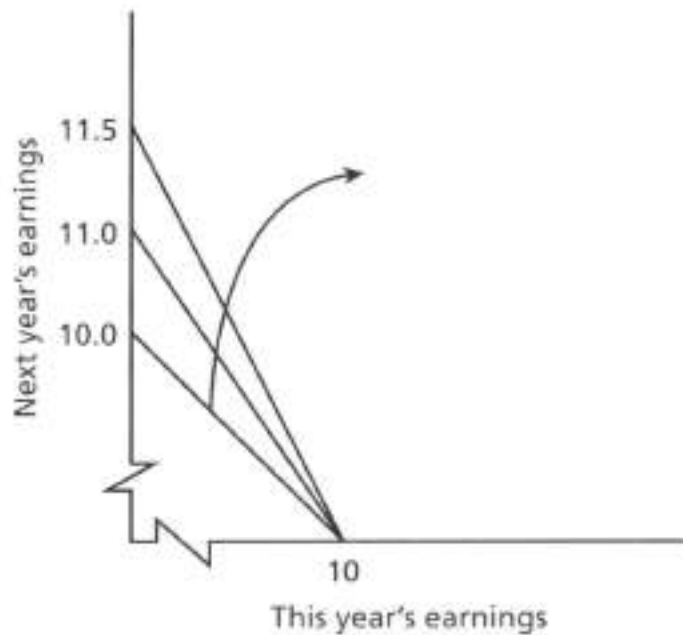
$$P_r = \frac{P_n}{(1+r)^n} \quad (14.3)$$

Going back to the example of the millionaires' club, we can see that if the discounting factor is 10 percent, then \$1 million received thirty-five years from now will be equal to

$$\frac{\$1,000,000}{(1+0.1)^{35}} = \$35,584.10$$

Alas, in light of this analysis, it appears that the dream of \$1 million has to be curtailed; the *present value* of \$1 million received thirty-five years in the future is only about \$35,000. With this amount, one certainly cannot expect all the trappings required for membership in a millionaires' club. This process is called **present value analysis**, by which we calculate the current or present value of a dollar to be gained in the future. We can use this formula to calculate the present value of a stream of benefits and costs to arrive at the net present value of a project. Notice that the larger the discount rate, the lower the present value of future dollars. If we were to discount \$1 million at a 15 percent rate, we would arrive at the paltry sum of \$7,508.89 for the same time period.⁶ You can see that as we increase the discount rate, the future dollars look smaller and smaller. As a result, gains to be made in the future look increasingly less attractive; similarly, the prospect of losses in the distant future looks less ominous. Therefore, the discount rate captures the strength of the desire to have money now as opposed to sometime in the future. This desire is called **time preference**.

The relationship between time preference and the discounted future value of a dollar is shown in Figure 14.3. In this figure, we have plotted this year's earnings on the

FIGURE 14.3 Time Preference and Discount Rate

horizontal axis and next year's earnings on the vertical axis. If we do not have a time preference, we will be indifferent between \$10 today and \$10 next year. The line connecting \$10 on the two axes shows an **indifference map** (we are indifferent between, or prefer equally, any two points along this line) with no time preference. However, if we discount the future earnings at a 10 percent rate, then to be on the same utility plane, we must earn \$11 next year. If we have an even higher time preference, equal to a 15 percent rate of discount, unless we earn \$11.50 in the following year, we would prefer to have \$10 today.

Let us consider a concrete example. Suppose we are evaluating two projects with the streams of benefits and costs shown in Table 14.2. From this table, it is clear that project A provides us with double the net benefit (\$80) provided by project B (\$40).

TABLE 14.2 Comparison of Costs and Benefits for Projects A and B

Year	Project A		Project B	
	Benefits	Costs	Benefits	Costs
0	0	30	15	5
1	0	15	15	5
2	0	10	15	5
3	10	5	15	10
4	20	5	15	10
5	120	5	15	15
Total	150	70	90	50

Should we automatically choose project A over project B? If we were to jump to this conclusion, we would be remiss, since we would fail to consider that the increased net benefits of project A come at a later stage in the project's life. Our choice between the two projects will depend on the strength of our time preference—the willingness to wait for future returns. Therefore, to compare the two projects on level ground, we must translate these future streams of benefits and costs into their present values. The formula is written as

$$PV = \sum_{t=0}^n \frac{(B_t - C_t)}{(1+r)^t} \quad (14.4)$$

where PV is the present value of the project, B_t is the benefit, and C_t is the cost at time t .

We can expand this expression and write

$$PV = \frac{(B_0 - C_0)}{(1+r)^0} + \frac{(B_1 - C_1)}{(1+r)^1} + \frac{(B_2 - C_2)}{(1+r)^2} + \frac{(B_3 - C_3)}{(1+r)^3} + \dots + \frac{(B_n - C_n)}{(1+r)^n}$$

Since any number raised to the power 0 is equal to 1, the expression can be written as

$$PV = (B_0 - C_0) + \frac{(B_1 - C_1)}{(1+r)^1} + \frac{(B_2 - C_2)}{(1+r)^2} + \frac{(B_3 - C_3)}{(1+r)^3} + \dots + \frac{(B_n - C_n)}{(1+r)^n}$$

You may notice that the 0th year's net benefits are not discounted. This makes eminent sense, because the current year's dollar is equal to its face value and hence does not need to be discounted.

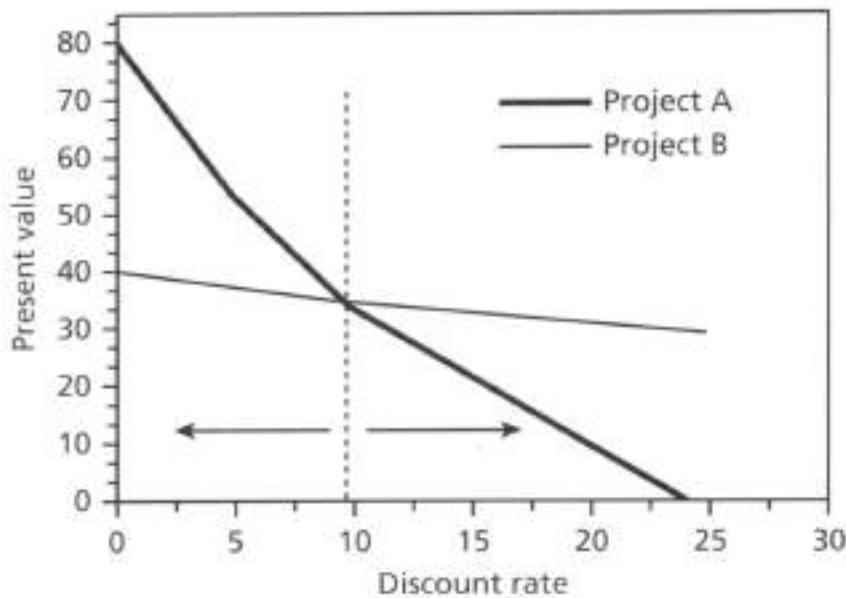
If we are willing to wait (that is, it does not matter to us whether we receive our payments today or tomorrow), our time preference is said to be nil. In such a situation, we discount the future stream of net benefits with a zero discount rate and therefore do not discount at all. Thus, if we do not have any time preference, the net benefits for projects A and B are \$80 and \$40. In contrast, suppose we do have a definite time preference, and we want to evaluate the future stream of net benefits for the two projects at a 10 percent discount rate. In such a case, we can write the present values of project A (PV_A) and project B (PV_B), discounted at 10 percent, as⁷

$$\begin{aligned} PV_A &= (0 - 30) + \frac{0 - 15}{(1 + 0.1)} + \frac{0 - 10}{(1 + 0.1)^2} + \frac{10 - 5}{(1 + 0.1)^3} + \frac{20 - 5}{(1 + 0.1)^4} + \frac{120 - 5}{(1 + 0.1)^5} \\ &= -30 - 13.64 - 8.26 + 3.76 + 10.25 + 71.41 = 33.52 \end{aligned}$$

and

$$\begin{aligned} PV_B &= (15 - 5) + \frac{15 - 5}{(1 + 0.1)} + \frac{15 - 5}{(1 + 0.1)^2} + \frac{15 - 10}{(1 + 0.1)^3} + \frac{15 - 10}{(1 + 0.1)^4} + \frac{15 - 15}{(1 + 0.1)^5} \\ &= 10 + 9.09 + 8.26 + 3.76 + 3.41 + 0 = 34.52 \end{aligned}$$

From the preceding calculation, we can see that discounted at a 10 percent rate, project A is less preferable than project B because it carries a lower present value.

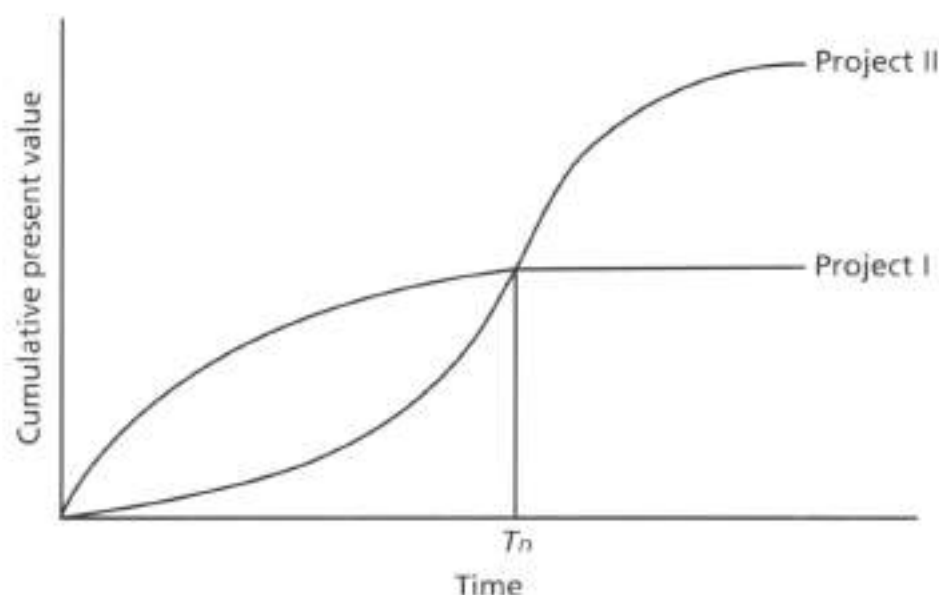
FIGURE 14.4 Plot of Present Values as a Function of Discount Rates

We can also see that the present values of the two projects will depend on the rate of discount, which determines their relative desirability. I have plotted the present values as functions of the rate of discount in Figure 14.4. From this figure, it can be seen that the two projects become equally desirable at a discount rate slightly less than 10 percent. For discount rates below 10 percent, project A is preferable to project B, but the relative desirability changes for discount rates of 10 percent and above. This change reflects how project A's benefits come at the end of its life. In contrast, project B yields positive net benefits from its inception. Therefore, if we can afford to wait (and have a small discount rate), we would prefer project A. However, if we are in a hurry to get back the returns on investment (and therefore have a stronger time preference), we should choose project B.

Finally, if the present value is negative (as project A is for a discount rate close to 24 percent), we should reject the project.

Choice of Time Horizon

The choice of an appropriate **time horizon** is of crucial importance for a cost-benefit analysis. The relative desirability of a project is intrinsically connected to when it ends. The length of a time period affects the desirability of long- and short-term projects. Many projects for drug interdiction are designed for short-term results. In these projects, drug use is considered primarily as a law-and-order problem, and efforts are made to lower the supply of illicit drugs by police action. In contrast, programs treating drug abuse are seen as long-term public health initiatives that require expenditures on education, rehabilitation, and employment opportunities. These demand-side efforts (trying to reduce the drug demand) offer longer term solutions. So unless they are given longer time horizons, their impacts on drug use will not be fully realized. Project desirability dependent on the choice of time horizon is shown in Figure 14.5. You can see that before

FIGURE 14.5 The Effects of Choosing a Time Horizon

the critical point in time, T_n , the short-term project (Project I) yields higher net present value. However, beyond this point, the long-term project (Project II) becomes more attractive.

Choice of Discount Rate

The results shown in Figure 14.4 demonstrate how sensitive the assessment of a project's desirability can be to a change in the discount rate. Therefore, it is essential that we come up with the "correct" discount rate in evaluating a public project. In fact, the sensitivity of a public investment decision to the choice of discount rate was underscored by a proposed joint water project between Canada and the United States. Whereas the United States analysts, who used a lower discount rate, recommended the project, their Canadian counterparts, who used a higher discount rate, rejected it. There is little mention in the cost-benefit literature of how to choose the "appropriate" discount rate. A simple personal example may explain one of the reasons for this confusion. Suppose I am thinking of investing in a project that yields a certain amount of money over a period of time. In considering the desirability of this investment, I may consider alternative ways to invest the money. I find out that the best return on available investment is 8 percent. In such a case, I would discount the future net benefits of the project at an 8 percent rate. If this discounting provides me with a positive net present value, I should invest; otherwise, I should not. This process of considering alternative uses of money provides the opportunity cost, or the **shadow price**, of money. From the standpoint of a local government, the alternative use of money determines its opportunity cost. That is, if we did not invest in the convention center, we could use the money for some other projects important to the residents. Without full information on the opportunity cost of money, we may not be able to determine the appropriate discount rate.

However, the opportunity cost of money may not be the only guiding principle for investment. I may choose to discount my investment with a rate that reflects my own time preference. People may use many different rates of discount representing different time preferences. Thus, if I want my money right now (or, in other words, I have a strong time preference), I would use a very high discount rate. If, however, I have a long-term perspective, I would be willing to wait for a higher return in the future, in which case the rate of discount would be quite low. In fact, in an extreme case, my time preference could even be less than zero. Imagine that you are the dictator of a small but wealthy country. You have all the money you want for the present. However, what you do not have is security for the future—you may find yourself deposed in a coup or a revolution. In such a case, your discount rate could even be negative, and hence you would be willing to put your money in a Swiss or otherwise secret bank account, whereby the bank would charge you for the safekeeping of your (presumably ill-gotten) assets and would not pay you interest. In other words, depending on my circumstances, it may be perfectly reasonable for me to use a discount rate different from the opportunity cost, in conformity with my personal time preference.

Similarly, in the case of society, economists have argued back and forth regarding the optimum rate of discount on the basis of opportunity cost and **social time preference**. The proponents of opportunity cost for investment argue that government pays for its investments by taking money away from private citizens; thus, unless the returns of these investments are at least equal to those of the private sector, it does not make any economic sense for the government to invest. According to this point of view, when evaluating a public project, an analyst should take into account the opportunity of alternative investment opportunity in the private sector and hence discount the project by a rate equal to the private rate of return. Suppose that NASA is proposing building a new kind of space vehicle, which would be able to deploy satellites with a greater degree of efficiency than do the existing space shuttles. Money for this project must be raised from private taxpayers. If the market is yielding 10 cents on a \$1 investment, unless it could be shown that the return from this project would be at least 10 percent (that is, the project carries a positive net present value when discounted at 10 percent), the project should not be undertaken.

This rule of discounting seems reasonable. However, upon further consideration, it turns out to be unsatisfactory. Problems arise because like all other markets, the capital market is characterized by various kinds of imperfections. For example, monopolies can significantly distort security prices.⁸ Also, the government may face several political constraints in its investment decisions, which can make any comparison with the private sector invalid. Therefore, by looking at the vast and imperfect capital market, an analyst is likely to be confused about the “appropriate” rate of yield in the private sector.

The comparison of government rate of return with that in the private market is further complicated by the often ill-defined source of government revenue. Thus, if tax money comes out of taxpayers’ savings, the consideration of opportunity cost may be valid. However, if it comes from the money that was allocated by taxpayers for consumption, the comparison would not make much sense. In addition, the sources of government revenue are diverse. The government may get its money from taxes, tariffs, licenses and fees, or selling bonds or government property. Not all of these moneys have

the same opportunity cost, and it is impossible to pinpoint the exact source of the funding for a project.

Finally, the outputs of government projects may be intangibles and as such cannot be measured in monetary terms. Thus, it would be impossible to convert the benefits of subsidized school lunch programs into strict monetary units. Also, many government programs can generate long-term positive externalities, which are extremely difficult to measure. For example, the development of computer technology, to a large extent, has been a by-product of the space program. Yet at the time of the inception of the program, nobody could have predicted this fortuitous outcome.

The recent literature in economics suggests the use of a discount rate that reflects the subjective time preference of society.⁹ Unfortunately, the scholars who have spent a great deal of time contemplating the appropriate discount rate are unable to tell us exactly which number to use.¹⁰ However, the insights derived from these discussions can at least point out the folly of using grossly inappropriate rates. Given the enormous complexity of our world, we may consider that benefit to be a giant step forward. In any case, the confusion about the correct discount rate is reflected in the practices of the highest federal agencies.

At the federal level, the Office of Management and Budget (OMB), the U.S. Government Accountability Office (GAO), and the Congressional Budget Office conduct discounting on a regular basis for capital expenditure programs, lease-purchase decisions, regulatory reviews, and sales of government assets. However, they use different discount rates.¹¹

For example, the OMB, which determines rates for all executive agencies, typically uses a 10 percent discount rate, corrected for the rate of inflation, for all tax-financed projects. If an agency wants to use a different rate, it must justify its choice. Since lease-purchase projects are funded by government bond financing, they are evaluated at the Treasury Department borrowing rate with a maturity date corresponding to the project's completion, plus 0.125 percent to cover the Federal Financing Bank's borrowing charge. Thus, a fifteen-year project will be discounted at the fifteen-year Treasury bond rate (say, currently at 4.75 percent) plus 0.125, or at 4.875 percent. In contrast, when a government asset is sold to nongovernmental entities, the OMB uses comparable private sector borrowing rates.

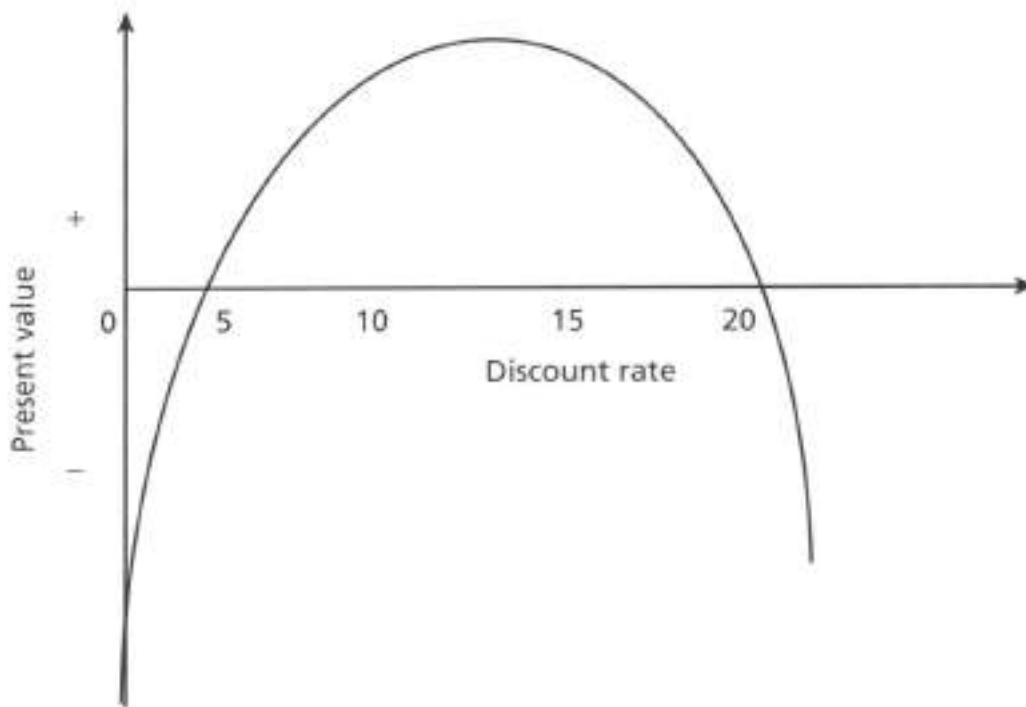
The GAO's use of a discount rate is based on the current yield of Treasury debt between one year and the life of the project. The GAO uses sensitivity analysis through different discount rates (see the discussion of sensitivity analysis later in the chapter).

The Congressional Budget Office, like the OMB, uses an inflation-adjusted Treasury borrowing rate and then looks at the sensitivity of the project by adding or subtracting 2 percent from the chosen discount rate. Also, like the OMB, it uses private sector rates to evaluate asset sales.

Because there is no unanimity at the federal agency level, state budget offices either follow the federal directives or use different rates. Local governments usually follow state guidelines.

The Internal Rate of Return

Since it is so difficult to settle on a universally acceptable discount rate for public projects, a decision maker can often be tempted to use what is known as the **internal rate of return**

FIGURE 14.6 The Internal Rate of Return

for judging the desirability of a project. The internal rate of return is defined as *the rate of discount at which the present value of a project is equal to zero*. Thus, from Figure 14.4, you can see that the present value for the hypothetical project A approaches zero at about 24 percent, which is its internal rate of return.¹²

The advantage of using an internal rate of return is that it reduces a decision maker's burden of having to make a choice on the basis of a single discount rate. Instead, a project can be accepted if the returns are larger than what can be reasonably obtained in an alternative investment. Going back to our example, if the internal rate of return for a project is 24 percent, it is so much higher than what can be reasonably expected in other possible investments that a decision maker will be hard-pressed to reject it.

However, despite this intuitive appeal, the internal rate of return suffers from some important shortcomings. First, although the internal rate of return is useful in pointing out the desirability of a single project, it is ineffective when choosing between two or more. Second, depending on the configuration of streams of net benefits, projects can often have more than one internal rate of return, as shown in Figure 14.6.

Since there are two rates (5 percent and 20 percent) at which the present value of the project becomes equal to zero, our decision maker is likely to remain confused about the desirability of the project. Edith Stokey and Richard Zeckhauser note that the internal rate of return can point to the correct public policy decisions under some rather unrealistic circumstances, such as when there are no budgetary constraints, there are no comparisons with alternative projects, and the stream of returns is first negative and then positive.¹³

CHOOSING THE BEST ALTERNATIVE

Sensitivity Analysis

After conducting an analysis, you may decide that you do not agree with your finding, or other experts in the field may disagree with you, or you may simply want to know how robust your finding is. In such cases, you may want to conduct a sensitivity analysis. This process allows analysts to change the assumptions underlying their analyses, estimate a different set of numbers, choose an alternative discount rate, and then see if there is a significant change in the outcome. Where there is a nonexistent market, you may impute monetary values on the basis of a different set of criteria. Or you may look into alternative scenarios for estimating future streams of income. For example, you may consider the most plausible case, the most optimistic case, and the most pessimistic case, with their corresponding subjective probabilities. If the probability of a disastrous case scenario poses an unacceptable amount of risk, that risk should be made clear to the decision makers. Finally, the use of different discount rates can also be of tremendous help in understanding the strength of a particular project. In our hypothetical case shown in Figure 14.4, you can demonstrate that project A is preferable up to a 10 percent discount rate, beyond which project B becomes preferable. Such a demonstration of sensitivity strengthens the robustness of your argument more than choosing a single rate and then drawing a conclusion on the basis of it.

Cost-Benefit Ratio

When choosing among alternatives, you may want to consider the ratios between benefits and costs, instead of their difference. In many cases, the use of either of these two criteria will lead to the same conclusion. However, a difference in the scale of operation may bring about a conflict. Consider the problem shown in Table 14.3.

From this table, you can see that the sheer difference in the size of the two projects means that we are really comparing apples and oranges. Therefore, using the ratio method, we are apt to choose project A, whereas using the difference method, we would have chosen project B.

The cost-benefit ratio, however, is often thought of as the efficiency measure. It may be particularly useful when you are evaluating a set of projects with a fixed amount of funds available for them. Let us consider the following hypothetical example. Suppose your city is considering spending \$10 million, raised through bond financing, to expand the public library system. Eight facilities are being proposed: one large facility at the center of the city and seven smaller facilities spread through various neighborhoods. Each one has a different cost estimate. The purpose of a public library is making its facilities accessible to the residents of the city, so you have decided to use the projected circulation from each facility as the measure of each facility's benefit. Table 14.4 shows the various benefits and costs.¹⁴

TABLE 14.3

Contradictory Decision Based on Net Difference and Benefit-Cost Ratio

	Benefit	Cost	Net benefit	Benefit-cost ratio
Project A	10	3	17	3.33
Project B	50	25	125	2.00

TABLE 14.4 Costs and Benefits of Library Facilities

Location	Benefits (projected circulation per year, $\times 1$ million)	Costs ($\times \$1$ million)	Benefit-cost ratio
City center	15.00	7.50	2.000
Neighborhood A	3.00	0.75	4.000
Neighborhood B	1.75	0.50	3.500
Neighborhood C	0.95	0.76	1.250
Neighborhood D	2.10	0.66	3.182
Neighborhood E	6.50	1.78	3.652
Neighborhood F	8.30	5.10	1.627
Neighborhood G	3.90	1.20	3.250

If you add up the costs of all eight facilities, you will quickly find that the project would cost \$18.25 million, far beyond the city's allocation of funds. Therefore, you may have to choose among these eight possibilities. In this case (Table 14.5), you may rearrange Table 14.4 according to the relative rates of efficiency, the cost-benefit ratio.

In Table 14.5, the column on cumulative cost tells us how much each facility would add to the total cost. Alas, reality often does not make the life of an analyst easy. In this case, listing the sites on the basis of their efficiency shows that the top five require \$4.89 million. Adding the city center facility would mean going over the \$10 million budget limit by \$2.39 million. Therefore, you may decide to drop this one and include the next efficient site, neighborhood E. In that case, the total expenditure comes to \$9.99 million, serving 25.55 million readers.

As you can see, the cost-benefit ratio served as a guide for our choice of the most efficient allocation of public funds. However, efficiency may not be the only consideration. For instance, the mayor may want the grandiose facility at the city's center for political reasons, or a more deprived neighborhood may be more deserving even though it does not score well on the measure of relative efficiency.

TABLE 14.5 Costs and Benefits of Library Facilities, Ranked According to Benefit-Cost Ratio

Location	Benefits (projected circulation per year, $\times 1$ million)	Costs ($\times \$1$ million)	Benefit-cost ratio	Cumulative cost ($\times \$1$ million)
Neighborhood A	3.00	0.75	4.000	0.75
Neighborhood E	6.50	1.78	3.652	2.53
Neighborhood B	1.75	0.50	3.500	3.03
Neighborhood G	3.90	1.20	3.250	4.23
Neighborhood D	2.10	0.66	3.182	4.89
City Center	15.00	7.50	2.000	12.39
Neighborhood F	8.30	5.10	1.627	17.49
Neighborhood C	0.95	0.76	1.250	18.25

The Limits of Cost-Benefit Analysis: Redistribution of Income

The preceding discussion of cost-benefit analysis leaves out one important problem: every government project creates winners and losers. How do you balance the gains against the losses? In a zero-sum society one group of individuals almost always comes out ahead of others. It is relatively easy to show the benefits of massive highway construction in the United States. This was especially true during the 1950s, when the increased ease of transportation gave a tremendous boost to trade and commerce. Yet such construction did not come without a long list of those who were adversely affected by this development. Many small towns by the now abandoned but previously well-traversed roads were simply wiped out of existence; many communities lost their identities as the highways caused dislocation, isolation, and a general deterioration of neighborhoods. The trucking industry gained at the expense of a more energy-efficient railroad industry; suburbs gained at the expense of downtown areas.

Yet as a reflection of the state of the art in social sciences and social philosophy, cost-benefit analysis is singularly unable to solve the problem of inevitably having winners and losers. I discussed the problem in chapter 4. Social philosophers have been grappling with the problem of determining distributive justice for centuries, with surprisingly few solutions. Paraphrasing Sir Winston Churchill, rarely in the entire field of social philosophy and social science have so many spent so much effort to produce so few results. The inability to address the question of redistribution not only remains the biggest challenge for cost-benefit analysis but also truly exposes the intellectual poverty of our social philosophy.

Cost-Effectiveness Analysis

The basic principle of cost-benefit analysis is simple enough: choose the alternative that gives you the maximum net benefit. However, in many cases of public expenditure, the benefit schedule is ill defined or impossible to measure: the local fire department wants to buy fire engines that can accomplish more or less the same task as the current trucks; an organization is planning to purchase personal computers of similar capabilities to the ones now in use. In such cases, if it is reasonable to hold that all the alternatives are substitutes, then we can assume that the benefit levels for all of them are equal and, hence, can be regarded as constant. Thus, we can choose the least expensive alternative.

Suppose, however, that money has been allocated to support an outreach program to teach children of homeless families, and proposals are being evaluated for choosing the best alternative. In this case, all the alternatives will spend the full allocation, so the costs are the same. For our comparison we need concern ourselves only with maximizing the benefits. In this circumstance we can use a truncated version of cost-benefit analysis. This is called **cost-effectiveness analysis**. The two simple rules for cost-effectiveness analysis are as follows:

1. If the benefits are the same, choose the alternative with the least cost.
2. If the costs are the same, choose the alternative with the most benefits.

Returning to the fire department example, two different makes of fire engines are under consideration. Both serve the same function, so we can concentrate on various aspects of costs of purchase and operation. Having done so, we can choose the one that is less expensive. Although the process seems simple enough, all the problems of a full-fledged cost-benefit analysis remain just as relevant. For a real-world application of the concepts discussed in this chapter, see "A Case in Point."



A CASE IN POINT

Fitting Seat Belts in Texas School Buses¹⁵

The problem: In 1985 Texas was confounded by the problem that more adults and children were injured in school bus accidents (a total of 635 in 1985) than in any of the previous eight years. Since Texas did not require school buses to be equipped with seat belts, only a very small fraction of the injured children (1.7 percent) were wearing them.

The goal: Evaluate the potential reduction in serious injury and death to the schoolchildren resulting from the installation of safety belts.

The alternatives: Install seat belts on all the school buses, or do nothing.

Assessment of costs and benefits: The benefit was the avoidance of serious injuries to children, and the cost was the expense involved in refurbishing the existing school buses. Imputing tangible values for intangible benefits was a bit trickier because data on the effectiveness of seat belts in the prevention of serious accidents were not available for school buses. Therefore, it was assumed that seat belt effectiveness in buses was the same as in automobiles, for which data were available. These data were used to calculate the percentage of preventable injuries. The severity of injuries was calculated using the Multiple Abbreviated Injury Scale (MAIS). The results are shown in the table.

Types of Injuries and Preventable Numbers for Texas, 1983–1985

Age and injury category	Number of accidents		Preventable fraction	Number of preventable cases
	Belted	Not belted		
MAIS 0: no injury	4	354		
5–14 years	3	297	0.00	0
15–18 years	1	57	0.00	0
MAIS 1: minor injury	3	253		
5–14 years	2	212	0.00	0
15–18 years	1	41	0.00	0
MAIS 2: moderate injury	1	150		
5–14 years	1	126	0.01	1
15–18 years	0	24	0.23	6
MAIS 3: serious injury	0	91		
5–14 years	0	76	0.07	5
15–18 years	0	15	0.39	6
MAIS 4: severe injury	0	37		
5–14 years	0	31	0.20	6
15–18 years	0	6	0.64	4
MAIS 5: critical injury	0	32		
5–14 years	0	18	0.25	5
15–18 years	0	4	0.71	3
MAIS 6: fatal injury (no fatal injuries were reported)				

Discount rate: A social discount rate of 6 percent was used to calculate the present value of net benefits over time.

Recommendations: The results indicate that the economic benefits from mandatory seat belts would not be cost effective for all the existing buses. However, mandatory seat belts could be cost effective for new buses.

Conclusion: This result is surprising given the widely accepted safety benefits resulting from automobile seat belts. The authors explain this discrepancy by noting that the number of injuries per mile of school bus ride is significantly less than for automobiles. Also, the severity of injuries is much higher in autos. Second, since the injury victims are children, who do not start earning income for years (in contrast to income-earning adults), the use of their discounted earning capability reduces the size of the benefits. Third, the indirect benefit that the habit of wearing seat belts may carry over to private automobiles, as the children learn to use them, was not included in the study. Finally, the study did not consider the cost-effectiveness of requiring seat belts for only the newly acquired buses.

Other considerations: School buses still transport millions of children every year, and every year there are a few accidents that cause death and injury to some children. After each accident, renewed calls are heard for installing seatbelts. However, the National Highway Traffic Safety Administration (NHTSA) claims that current school buses are among the safest forms of transportation. After studying the results of crashes involving school buses, the NHTSA recommended against federally mandated retrofitting of the old buses or installing seat belts on new ones. Since most accidents involving school buses were either frontal or rear collisions rather than side impacts, the NHTSA concluded that because of the length of the buses, the children were amply protected by a safety feature called compartmentalization. This involves seating passengers in rows of padded seats with cushioned backs, so that in the event of a collision, passengers are either pushed back into their seats or thrown forward into the padded backs of the seats ahead.

The NHTSA also believes that belts on school buses may hamper rescue or evacuation efforts, as dazed children, unlike adults, may have difficulty unbuckling themselves. Some mischievous students may also use the heavy buckles as makeshift weapons during fights, creating even more of a safety hazard.

Finally, if seat belt use is made mandatory, there can be the additional problem of compliance. Since bus drivers already have a significant amount of responsibilities, schools might have to hire additional monitors to ride on all the buses. All of these can open up new allegations of negligence and numerous legal actions, which cash-strapped school districts are unable or unwilling to bear.

Discussion Points

1. How was cost-benefit analysis used in determining whether to fit seat belts on existing buses?
2. Do you agree with the study? If you were doing the analysis, would you have altered any of the assumptions? How do you think that might have changed the recommendation?

KEY WORDS

- | | |
|--|---------------------------------|
| Consumer surplus (p. 346) | Negative externalities (p. 347) |
| Cost-benefit analysis (p. 345) | Net social benefit (p. 351) |
| Cost-effectiveness analysis (p. 366) | Pecuniary effect (p. 351) |
| Direct costs and benefits (p. 348) | Positive externalities (p. 347) |
| Discount rate (p. 356) | Present value analysis (p. 356) |
| Discounted future earnings (p. 352) | Required compensation (p. 352) |
| Face value of life insurance (p. 352) | Shadow price (p. 360) |
| Indifference map (p. 357) | Social time preference (p. 361) |
| Indirect costs and benefits (p. 348) | Time horizon (p. 359) |
| Intangible costs and benefits (p. 348) | Time preference (p. 356) |
| Internal rate of return (p. 362) | |

EXERCISES

- Suppose you are considering buying a new home or renting a bigger apartment. In a report, define your goal, the feasible set of alternatives, and the set of desirable attributes. Then, given your budgetary limitations, write a report describing the process by which you arrived at the optimum choice.
- Consider a project in your city that proposes to increase the supply of any public good. Indicate the benefits and costs, pointing out the possible area of consumer surplus, the pecuniary effect, and the loss of producer surplus. You may also think of a project that proposes some new government regulation to bring social cost in line with private cost. Point out the area of possible dead-weight loss in consumer surplus, the pecuniary effect, and the gain in producer surplus. Explain your choice of discount rate (you do not need to show any actual calculation of present value).
- What is a social discount rate? Why is there so much confusion in defining the term *social discount rate*? Why is it important to determine this rate when calculating the relative desirability of a public project?
- What is a cost-effectiveness study? How does it differ from a full cost-benefit analysis? Give examples of when it is more appropriate to use a cost-effectiveness analysis.
- Calculate the net social benefit of the two projects in the table by evaluating them at 0, 5, 10, and 20 percent discount rates.

Year	Project A		Project B	
	Benefits	Costs	Benefits	Costs
0	10	10	0	55
1	15	10	5	25
2	20	10	15	25
3	30	10	50	15
4	20	10	65	15
5	10	10	95	15

6. Suppose your final answer in a popular game show was correct and you received \$1 million. The show offers you two options: you can take your \$1 million in a lump sum, or you can receive \$145,000 per year for the next twenty-two years. Within the framework of the present value analysis, which option would you choose and why? Explain your choice of discount rate. One more small matter of information: Regardless of when you take your prize money, you are subject to a tax of 40 percent.
7. A recent study has demonstrated the link between death rates and the amount of fine particulates less than ten microns (one-thousandth of a millimeter) across.¹⁶ These particulates come from cars, trucks, power plants, construction, and even agriculture. The study found that for each ten micrograms of particles per cubic meter of air over a twenty-four-hour period, the death rate from all causes rose more than 0.5 percent. That is, in a large city in which 100 people die each day, a rise in the particulate pollution of twenty micrograms per cubic meter can cause an additional death per day. During the study period, New York averaged 190.9 deaths per day. The corresponding figures for Los Angeles and Chicago were 148 and 113.9.

Suppose you are an EPA analyst in charge of producing a position paper on a new pollution standard reflecting the new findings. How would you develop your study? Specifically, what kinds of information would you seek to demonstrate the desirability of your proposal through a cost-benefit analysis?

Notes

1. This line of reasoning has a long history and a voluminous literature in economics. For a seminal explanation, see Gary S. Becker, *The Economic Approach to Human Behavior* (Chicago: University of Chicago Press, 1976).
2. See Shoshana Grossbard-Shechtman, *On the Economics of Marriage: A Theory of Marriage, Labor, and Divorce* (Boulder, Colo.: Westview, 1993), esp. 5–84.
3. For an excellent discussion of the required compensation principle, see Edward M. Gramlich, *A Guide to Benefit-Cost Analysis*, 2nd ed. (Englewood Cliffs, N.J.: Prentice Hall, 1990), 67–70.
4. Kirk Johnson, "G.E. Facing Order to Remove Poisons from the Hudson," *New York Times*, December 6, 2000.
5. "Secret Studies Put Spill Damage at \$15 Billion," *Los Angeles Times*, October 8, 1991.
6. You can calculate the present value by using a simple hand calculator. If you have a y^x button on your calculator, you can make the necessary calculation. To calculate the present value of \$1 million thirty-five years from now at a discount rate of, say, 5 percent, punch in the numbers in the following sequence: $1,000,000 / (1.05^{y^x 35}) =$. This sequence will give you the answer: \$181,290.29.
7. You can calculate the net present value using Microsoft Excel. However, if you do, you should know that Excel assumes that discounting starts at period 0. Therefore, to obtain the results in this book, you should start discounting from the second year on and then add the net present value of the initial year.
8. Recall the security exchange fraud during the late 1980s and early 1990s by the giant trading houses such as Drexel Burnham Lambert and Solomon Brothers, which had a significant impact on the market. Similar irregularities in Japan caused widespread concern over the integrity of Japanese financial institutions.
9. See David F. Bradford, "The Choice of Discount Rate for Government Investments," in *Public Expenditure and Policy Analysis*, 3rd ed., ed. Robert H. Haveman and Julius Margolis (Boston: Houghton Mifflin, 1983), 129–144.

10. Edward Gramlich suggested that the social time preference for the federal government, corrected for inflation, in 1988 was about 4 percent for projects financed by tax revenue and 6 percent for those financed by bonds. See Gramlich, *A Guide to Benefit-Cost Analysis*.
11. John L. Mikesell, *Fiscal Administration: Analysis and Applications for the Public Sector*, 5th ed. (Fort Worth, Tex.: Harcourt Brace, 1999), 253.
12. Using Microsoft Excel, you can calculate the internal rate of return. For the example of project A in Figure 14.4, the internal rate of return is about 23.7 percent.
13. Edith Stokey and Richard Zeckhauser, *A Primer for Policy Analysis* (New York: W. W. Norton, 1978), 167.
14. You may notice here that although I have defined benefits in strict monetary terms, in this case we are using a nonmonetary measure. When a monetary measure is not feasible, and everyone can agree upon a readily available indicator of benefit, there is no harm in using such a nonmonetary index.
15. Adapted from Charles E. Bagley and Andrea K. Biddle, "Cost-Benefit Analysis of Safety Belts in Texas School Buses," *Public Health Reports* 103 (September–October 1988): 479–488.
16. Jonathan N. Samet, Francesca Dominici, Frank C. Curriero, Ivan Coursac, and Scott L. Zeger, "Fine Particulate Air Pollution and Mortality in 20 U.S. Cities, 1987–1994," *New England Journal of Medicine* 343 (December 14, 2000): 1742–1749.

So You Want to Be an Analyst? Some Practical Suggestions

In the wide-open field of public policy analysis, the term may conjure up images of a civil servant tucked away in the corner of a sterile office writing long reports on some specific areas of government operation. In today's world, policy analysts come in many different shapes and forms. They may work for the public sector, the private sector, or for-profit or nonprofit organizations. They may work for think tanks or advocacy groups. They may work for politicians or for law firms, for newspapers or radio or television stations, as part of teams or by themselves as private contractors. They may prepare long reports, opinion pieces, position papers, or policy memos. They may contribute to blogs, appear on talk shows, or make slide presentations for their clients. They may look at health policy, foreign policy, or budgetary policy. They also get involved in every level of government: local, state, federal, and international.

In this extremely varied arena, one commonality stands, however: analyzing public policy is more of an art than a science. That is, scientific "objectivity" can get you up to a certain point, but it gives you no guarantee of results. In this book, I have discussed many different tools of scientific research, from simple to relatively sophisticated. Through them we aim to develop "rational" or "scientific" or "objective" analyses. All of these terms imply strict impartiality, adherence to stringent methods of analysis, and an absolute absence of emotion. Yet recent advances in neuropsychology are showing that contrary to the accepted belief, our rationality cannot be separated from our emotions. After all, wisdom is about achieving a proper balance between the emotional and analytical parts of your brain. Unfortunately, nobody really knows where that proper balance might be. Therefore, in a subjective, uncertain, and intensely political world, I can offer you a few guidelines.

Basically, when approaching a problem analytically, we want answers to a stream of related questions: What do we know? What

TABLE 15.1 Steps in Conducting an Effective Analysis**Before you start*****"Know thyself" and your organization***

- Define your role.
- Do you know your client?
- Do you have a basic understanding of the issue?
- Remember that in public policy analysis, "objectivity" simply applies to the process.

Be aware of biases

- Do you think you already know the answer?
- Do you see a value conflict?
- Do you have a conflict of interest?
- Is there institutional pressure?

Be aware of external constraints

- Is there enough information?
- Is there enough staff, money, and time to do the analysis?

Begin your analysis***Organize your thoughts: a quick analysis***

- Think, decompose, simplify, specify, rethink.

Conducting a more elaborate analysis

- Define your goal.
- Gather information.
- Measure dependent and independent variables.
- Choose the right tool.
- Project the outcomes and make your recommendation.

Tell a good story: effective presentation

- Know your audience.
- Know what you don't know.
- Tell a compelling story.

don't we know? What do we need to know? What don't we need to know? What do we need to know to know what we do and don't need to know? However, let us go about asking these questions more systematically, in three sections: (1) before you start, (2) begin your analysis, and (3) tell a good story: effective presentation. I have placed the outline of my suggested steps in conducting an effective analysis in Table 15.1.

BEFORE YOU START**"Know Thyself" and Your Organization**

To "know thyself" is the first principle of western Hindu philosophy. According to ancient wisdom, it is the most difficult quest in life. Before starting an analysis, it may

be worthwhile to look into your own role as an analyst. As part of this query you must address the question, who is your client? For instance, suppose you are working as a staff analyst for one of the U.S. senators from your state, and you have been asked to prepare a position paper on welfare reform. In that case, you have to remember that you are serving a particular elected official from a specific political party. Even within a political party, members are associated with various wings, based on their ideologies. Therefore, unless it goes against your moral compass, one of the goals of your position paper may explicitly or implicitly recognize the senator's ideological orientation or one that the senator may be able to defend politically. When there is a strong moral dilemma, as we have seen in chapter 2, you may need to evaluate your position using a proper mix of exit, voice, and disloyalty within the organization. If, on the other hand, you are working for the neutral Congressional Budget Office, where presumably your client is the entire nation, and you prepare a report on the same topic, your orientation may differ.

When working on a specific issue, you should ask yourself, How much do I know about this? If you work as a staff analyst, you may need to get involved in all sorts of projects. Today, you are preparing a position paper for your senator on welfare reform; tomorrow, you may be asked to write one on a particular matter of military procurement. If you do not have much knowledge of the subject matter at hand, you must develop your own knowledge base through proper research.

In our analyses, we all seek objectivity. Yet as we have seen throughout the book—from the discussion of human nature to political orientation—we must deal with our subjective assessments every step of the way. In this confusion, we would do well to recall that “objectivity” of analysis should involve the process by which we find our solutions, not the solutions themselves.

Be Aware of Biases

Biases, as I have shown, are impossible to avoid. However, what is possible is to know the existence of your presuppositions. For instance, before you start your analysis, ask yourself if you think you already know what your conclusion will be. If you do, and if you are planning to conduct an unbiased analysis, it may be time for you to look at your own motivations. In that case, it will be important for you to get acquainted with opposing points of view. By synthesizing the various perspectives, you may arrive at an impartial assessment.

To be neutral is a worthy goal. However, sometimes it is impossible to avoid such conflicts when it comes to deeply emotional issues relating to religious faith, abortion, or gay rights. In case you do have a very strong view that you know to be immutable, and if your moral, ethical perspective stands against conducting an impartial analysis, you may try to recuse yourself from working on such a project. However, in many organizations, it is impossible to avoid such responsibilities. In such a situation, you may do well to explicitly state—either orally to your supervisor or as part of the written report—your own ideological position on the matter at hand. You should also be clear in your mind about any conflict of interest in the area you have been asked to investigate. Do you or any member of your family or circle of friends stand to gain (or lose) from the project? Sometimes these conflicts can have legal implications; at other times they may pose moral issues. In any case, it should be clear in your mind that there is no conflict of interest.

Finally, you must be aware of any organizational pressure that may be brought upon you to arrive at a certain conclusion. These pressures can be explicit or extremely subtle. Every organization, when examined closely, is built around a central belief structure. For instance, the U.S. Department of Veterans Affairs aims to help retired members of the military, the U.S. Department of Education generally attempts to aid students and schools, and U.S. Customs and Border Protection strives to keep illegal immigrants from coming into the country. Many think tanks promote certain ideologies, while advocacy groups often lean toward cherry-picking evidence to support their client groups or beliefs. In fact, in performing their duties, organizations often embrace a certain philosophical orientation. Therefore, any analysis that is contrary to its central article of faith may be considered heretical and can result in pressure being placed on the messenger. For example, supervisors may try to quash a report and stop its release, prevent a report from going up the chain of command, transfer an offending analyst to a different (often less desirable) position, or even try to fire an analyst. These are matters of explicit reprimand for the dissenting voice. Group pressure can also be subtle, from loss of friendship to being left out of happy hour. Psychologist Irving Janis popularized the concept of groupthink, whereby members of a group, often unconsciously, start thinking along a similar line.¹ Any deviation from this group orientation can result in the ostracism of the offender. Facing such group pressures, you may have to think of your options, among a combination of exit, voice, and disloyalty.

Be Aware of External Constraints

Apart from your own knowledge, experiences, and perceptions, and the organization you work for, there are certain external constraints that can affect your ability to properly analyze a problem. These constraints arise because, after all, life itself is limited. Professor Herbert Simon believed that life was about satisficing—reaching a sufficient result within a set of constraints—not about boundless optimization.

While conducting research on most topics, you may come across information limitations. I call it the tyranny of data.² Information is collected by people who may or may not have the same research agenda as you do. Therefore, quite often, what we can prove is dependent on what information is available. For example, if you want to show that students' scores on standardized tests are determined by the stability of their parents' marriages, you may find data on test scores and divorce rates, but the two variables may not correspond. Therefore, unless you want to develop a full-scale survey and collect your own primary data, there is no way for you to test your hypothesis. The problem, however, is that if you want to conduct a survey, it requires money, time, and specific expertise.

This brings us to the next constraints: money and time. The poet Homer famously wrote, "The gods are jealous of men because they die." Analysts, on the other hand, may be jealous of gods because they don't, and have unlimited time. Therefore, in designing your analysis, you must be realistic. You must learn how to cut your proverbial coat according to your cloth. This may seriously affect what you might consider the proper analysis. However, given the realities, you must learn to draw the line somewhere. Furthermore, not every problem merits elaborate analysis. Many are matters of mundane concern. Many require quick decisions. In such cases, you must learn to do a "back of the envelope" analysis on the basis of some accepted "rules of thumb," rather than

gathering primary data and using sophisticated statistical analyses. I argue, however, that whether you do a quick “back of the envelope” analysis or design a full-scale research project, the basic steps remain the same. The only difference is in the amounts of time and money that you and your organization commit.

On this point, I will add one more piece of guidance: policy analysts are trained at institutions of higher learning, but the expectations surrounding a policy paper may differ significantly from that of the academic environment where the analysts were educated. Since academics are often not as constrained by time and face a different set of expectations when they publish in scholarly journals, their analyses are often much more extensive. In comparison, a policy analyst’s report may appear shallow and superficial. However, we must remember that the two are writing for completely different clients; as a result, an analyst’s report rarely would be accepted in an academic journal, and an academic study may be quite inappropriate as a policy report.

BEGIN YOUR ANALYSIS

Organize Your Thoughts: A Quick Analysis

You have just been given the task of analyzing a particular policy issue. Where would you start? Behn and Vaupel have given us some excellent ideas.³ They propose the following five steps:

1. Think
2. Decompose
3. Simplify
4. Specify
5. Rethink

Think. The process of decision making starts with thinking. The thinking part essentially involves framing the problem. Because no problem in life is one-dimensional, you must look squarely at the heart of a complex, multidimensional problem. This reduction of a large problem into a manageable, bite-sized form often requires the use of simple numbers—numbers that can convey the essence of a problem. The percentage of people living in poverty and the comparative scores of graduating seniors on standardized tests from comparable countries are examples of such core numbers that clearly define a problem.

Decompose. Behn and Vaupel point out that the essence of analysis is decomposition: the ability to see a problem as an amalgam of various elements. Our brains offer us only limited space for processing complex information. Therefore, if we can see a problem in its various parts—some complementary and some contradictory—we have made significant progress in the right direction. Suppose you are writing a term paper on poverty in America. If you are fixated on the entire issue and go to the library and, finding hundreds of books and articles on the topic, get even more frustrated, you will not know how or where to begin. On the other hand, if you quickly choose a working title for your paper, which you can change as you go along, it will help you focus on a specific aspect of a very large problem. The second step for writing a term paper is to develop a table

of contents. This is the core of decomposition. Now you are looking at the problem in an organized fashion. Once you have decided on a title and developed a table of contents, you are on your way toward writing a focused paper.

Simplify. If we had unlimited time and money, we could do an exhaustive analysis. However, all of us must work within our means. Therefore, when time is of the essence, we do not have the opportunity to go through all the possible alternatives and outcomes to achieve the best solution. In such cases, we need to simplify the process by drastically eliminating all but the most essential components of the problem. The process of structured simplification is so effective in problem-solving situations because (1) without simplification, a problem can push us beyond our cognitive capabilities; (2) a simplified problem propels it from the realm of intuition or gut reaction to informed judgment; and (3) once a problem has been simplified and broken into essential components, decision makers' intuitions can be used to their fullest advantage. Finally, along this line, I should mention that in life there are people who are obsessed with perfection. There is of course nothing wrong with being a perfectionist, unless it prevents you from making a decision when you absolutely must. The simplification of a complex problem allows you to manage your time effectively. The value of the ability to deliver on time simply cannot be overstated within an organization.

Specify. When faced with a complex problem, after you have decomposed and simplified, try to draw a diagram (such as a decision tree, discussed in chapter 13) that specifies its interrelated components. Make a quick drawing of all the possible alternatives and options, including the option of doing nothing. Doing so will force you to specify the amount of risk and the amount of reward. You may put down your subjective assessment of risk with either probability figures or classifications, such as "highly probable," "probable," or "nearly impossible." Your pictorial rendition of the problem will give you deep insight into its true nature.

Rethink. On the basis of the quick decision rules, you have now arrived at a solution. Should you stick to it? Perhaps not. If there is time, rethink! Since all analysis is ultimately incomplete and is based on aspects of subjective judgment, you should not waste the opportunity to rethink. If there is enough time and you are still uncertain about your decision, start from the beginning once again, and make a second cut. Go on rethinking until you have come to the conclusion that given the constraint of time, you have done your best.

Conducting a More Elaborate Analysis

Define the Problem. Every analysis begins with a problem that needs a solution. It is important to recognize that your client must share the perception of this problem. Sometimes, in our activism, we may lose sight of the fact that our client may not perceive this to be a problem or believe that the solution is politically feasible. If there is a gap in perception, you must begin with identifying reasons it is in the interest of the decision makers to address this particular problem and end with an analysis of the political, economic, and organizational feasibility of the solutions you are proposing. How you define a problem may ultimately determine its fate. This will require many skills far beyond what is needed to conduct a technical analysis. In particular, your

definition of the problem must resonate with the organizational needs of your client. For instance, all elected officials see their constituents as their ultimate clients. Therefore, in proposing a project, you may appeal to their ideology or show that it would bring the greatest good to their home districts. Bardach calls this the “issue rhetoric.”⁴ The art of finding the right rhetoric is something that comes with a deep understanding of your client and the organization for which you are working.

Gather Information. Evidence is key to scientific inquiry. In today’s world, the ability to verify hypotheses with facts is as central as breathing oxygen. After you have defined the problem, you must actively engage in gathering facts. In the era of the Internet, a few chosen keywords can bring you a mother lode of information. Unfortunately, it is nearly impossible to have a drink at the mouth of a gusher. With so much information around, you must develop a strategy to gather information before you take a plunge.

Your strategy for gathering information must begin with knowing what you are looking for and what you are not. Your judgment about this must be backed by reason. For instance, suppose you are conducting research on poverty in America. Should you look into cases of comparable countries? At first glance, it may indeed sound like a good idea. However, your decision must be guided by all the external and internal constraints I mentioned above. How much time, money, staff, and expertise you have at your disposal should determine how widely you cast your net.

However, regardless of where you start and how far you want to take your analysis, you must do a quick review of existing books, articles, and reports relating to the problems at hand. In chapter 8, I detailed several sources of secondary information. Many sources and studies will not only provide you with invaluable insight into the causal linkage between your dependent and independent variables but may also offer examples of what happens when a particular policy is implemented. Their experience can allow you to prepare a “best practices” report.

As you look at the prior research and findings, you must be aware of the ideological underpinning of the studies. It will be up to you to determine if these studies’ orientations match with that of your own organization. As you go along, you must also consider the veracity of the sources of the information used in the studies.

Finally, in a rapidly changing world, where new technologies are creating brand new issues every single day, you may be hard-pressed to find a comparable case. When that happens, you may look into analogous examples. For instance, there are persistent allegations that photo-enforced intersections have turned into cash machines for local governments and the companies that install these devices. For instance, in 2008, Los Angeles earned more than \$3 million from these red-light cameras. As other cities scramble to install the devices, many are questioning the true motives of city officials. For instance in 2009, Kansas City not only began installing red-light cameras but also increased the fine for running a red light to \$100, while in other cities, fines are as high \$400 per infraction.⁵ The introduction of this new technology, much touted for its ability to reduce accidents, has created controversy, as some have suspected that the companies that install the cameras have reduced the time that yellow lights stay lit. While a shorter yellow light catches more offenders, studies of a number of cities (such as Mesa, Arizona, and Fort Collins, Colorado) show that a longer yellow light actually reduces the number of accidents, as drivers don’t speed up at the last moment to beat the camera.

A study by the Tennessee Center for Policy Research reports that an increase of 1 second reduced accidents by 73 percent, while a boost to 1.5 seconds reduced the frequency by 94 percent. Yet at the same time, the increase in the time the yellow light remained lit drastically reduced the revenue of the cities and the camera vendors. If you begin to write a report on red-light cameras for your own town, you may not find anything comparable in your area, but you may find plenty of analogous examples from other cities. However, in providing analogies, you must be careful to make sure they actually fit. That is, if you are basing your arguments for New York City on what happened in Mesa or Fort Collins, you must make sure that there are no intervening variables that would make your case null and void.

Measure Dependent and Independent Variables. Once you have reviewed the relevant literature, it is time to set your hypotheses. Calling a statement a “hypothesis” does not necessarily imply a prelude to statistical testing; rather, it simply means that you are making a conjecture that the variable you are targeting to change (the dependent variable) will be affected by altering a set of policy variables (the independent variables). Once you have identified the variables, you can proceed to test their relationships by using statistical tools or simply citing other examples. In the yellow-light example, you can either collect your own data by sending observers or using meters, or you can use the results obtained by other research institutions, academics, or government agencies.

We are inveterate producers of numbers. As I showed in chapter 6, we tend to express in numbers many things we see and even feel. Numbers have a magical quality of appearing precise and objective. However, as we have seen, like all things in life, social data are “contaminated” by subjective judgment. Therefore, once again, I should remind you that before you use any numbers for your dependent and independent variables, you must make sure you are measuring the right thing. For instance, when using data for accidents, you may assume that all data define an “accident” in the same way, but conclusions can be radically different when you look closer into the data and divide them into fender benders, accidents involving injuries, and those that cause fatalities. In each case, the paths of causal links with the independent variables can be significantly different. For example, you may find out that increasing the time of the yellow light reduces fender benders but actually increases the frequency of fatal collisions.

Identify Alternatives. After you have established a causal connection or gained an understanding of what is causing the problem at hand, you can focus your attention on how to achieve your goals of solving, mitigating, or managing it. These are the various policy options the government can use. In the course of conducting an analysis, this is the step that requires the greatest imagination and creativity. There may be many ways to skin a cat, but only a few are acceptable or feasible. While you may cast your net widely and think of many different alternatives, soon you will begin the process of pruning these options, as they are deemed too expensive, too time-consuming, too intrusive, too cumbersome, or politically infeasible. In other words, you will consider only a feasible set of options, given your constraints of time, money, public acceptance, and political feasibility.

However, you must recognize that all of these factors can change. What is considered to be too expensive today may bring money flowing to it under different circumstances.

For instance, after the 9/11 attacks, some raised the specter of terrorists' coming across the border from Mexico. Suddenly, a number of politicians began pushing for an impenetrable wall that would stretch from the Gulf of Mexico to the Pacific Ocean across the entire U.S.-Mexican border. As frenzied speeches raised the temperature and the nation reoriented its priorities, an enormous amount of money was allocated for the border fence. Similarly, as banking regulations weakened with the repeal of the Glass-Steagall Act of 1933 and the Securities and Exchange Commission became less aggressive in developing and enforcing regulations, the nation faced an economic crisis of a magnitude not seen since the Great Depression. The 2008 economic crisis suddenly altered the political climate, allowing both the Bush and Obama administrations to pass sweeping changes as the federal government took over some of the largest financial institutions. Suddenly, these intrusive government actions, which would have been considered infeasible months before the downturn, were accepted. Hence, while judging the feasibility of policy options, we must keep in mind that it is a relative concept that can change dramatically in response to large shifts in economic and political reality.

Identify Criteria. The most commonly used criterion for choosing the best course of action is efficiency, which is defined simply as the "most bang for your buck." However, as I discussed in chapter 1, in reality, monetary efficiency is not the sole criterion for choosing a public project. The drive toward economic efficiency must be weighed against concerns of equity, liberty, and security. Since these concepts are vague, most often the adoption of a criterion depends on how the political elites choose to frame an issue. The idea of global warming, which at one time was considered to be the quaint obsession of a handful of scientists and environmental extremists, has become mainstream over time. As a result, when we evaluate a public project, we not only use the criterion of dollars and cents; along with this, we often look at a project's "greenness." For instance, local government projects, apart from the mandatory completion of environmental impact statements, often tout their beneficial impact on the environment.

Choose the Right Tools. You have learned a good number of quantitative techniques, and you are ready to test your hypotheses or to project the outcomes of a proposed plan of action. Which model should you use? This is a very important question, which I have not addressed in these pages. Sometimes the problem at hand will dictate the type of tool you use. For instance, if there is no past information, you cannot use data-driven techniques such as regression analysis. However, you must keep in mind that every technique has its own limitations. There is nothing you can do about this except be fully aware of these limitations. If you are asked to defend your choice of tool, you should be prepared to do so.

Finally, I should mention that when choosing your tool of analysis, you must not be too enamored of the complexity of the tool itself. Many statistical analyses are extremely complex and fall beyond the scope of this book. Sometimes a researcher may be attracted by the apparent "sophistication" of a technique. You must remember that complexity, by itself, is no virtue. If there is something special in a technique that makes it particularly useful for the problem at hand, you should use it. Otherwise, given a choice, you should always choose the approach that is simpler to understand and to explain.

Project the Outcomes and Make Your Recommendation. Looking into the future is, at best, an inexact science. In this book I have discussed many different methods of prediction, each of which has its own strengths and weaknesses. Although public policy analyses require forecasting, for some outfits, forecasting is their entire business. For instance, the government forecasts future economic growth and inflation, hurricanes and other natural disasters, and the path of the latest epidemic. Therefore, before we get into a discussion of prediction, allow me to remind you of two rules that are worth remembering: (1) it is extremely difficult to forecast, especially the future, and (2) if forecast you must, forecast and forecast often. Although they are said in jest, let me explain why they give us valuable insights.

In forecasting the future, physicist Werner Heisenberg learned that it is of course impossible to fully penetrate the shroud of uncertainty (chapter 8). Therefore, in an open society, public policy forecasts cannot peer into the future without a good deal of uncertainty. Any claim of a guaranteed outcome is bound to be false.

Furthermore, forecasting in the real world must not be confused with oracular pronouncements, made only once and then expected to come true. Rather, similar to weather forecasting, we must continue to update our predictions as the objective conditions change. If we recall these two rules, we may learn to have realistic expectations regarding foretelling what is to come.

Despite these shortcomings, any policy discussion must contain an effort at forecasting. When there are no past data, we must rely on various subjective assessments, from forecasting by analogies to the use of more structured forms, such as the Delphi method. If we have time-series data, we can certainly look for a trend. However, no analysis of trend can predict a turning point. That is, if the economy is on the rise, we can predict where it will be at some point in the future, but if it is currently declining, regression models for projecting the trend may be inadequate to determine where the bottom might be. For those cases, we may use a multiple regression model, looking into the causes of a phenomenon (see Appendix B). Once again, given the fact that these models predict on the basis of an average, our forecasted values may not have a great deal of policy significance, although a causal model may provide us with the knowledge of what is caused by what.

When all manner of "scientific" forecasting fails us, we turn to another method: intuition, or simply a "hunch." Of late, a number of economists have been looking into intuition. Their research has revealed that when an expert is asked to forecast simply on the basis of intuition, the expert often performs better than even the most sophisticated models.⁶ After all, the complexity of the human brain is no match for even the most mathematically complicated data-driven statistical analyses. Therefore, while we develop our policies on the basis of forecasting future outcomes, we must learn to live with life's uncertainties. The only danger we need to be concerned about is the presence of systematic biases and wishful thinking (see chapter 10).

TELL A GOOD STORY: EFFECTIVE PRESENTATION

You have done an excellent analysis. Now comes the final hurdle: it is your turn to present it. You can present your findings along with your recommendation in the form of a written report or in front of a group of those whose job it is to make the actual decision.

They can be the president, members of the Congress, other elected officials, or, in the case of private corporations, the directors or your superiors. If you think your findings are the best analysis of the problem, your job now is to convince them of the efficacy of your work. You should know that doing the analysis and making an effective presentation require two different sets of skills. You may do an excellent study but may fail to persuade the decision makers. Therein lies the art of persuasion. Let us discuss some of its essential elements.

Know Your Audience

You must understand your audience. What may be effective in one setting can completely fall flat in front of another group. The groups may differ in their knowledge of the subject matter, their levels of sophistication, and their interest. For instance, in the days after the catastrophic 9/11 attacks, many military planners as well as political leaders in the United States lacked detailed knowledge of the Islamic world; in analyzing the links between various groups, many were unaware of the deep religious divisions that exists between Shi'a and Sunni Muslims. The 9/11 Commission hearings, conducted years later, revealed its members' ignorance of some of the most basic ground realities of the Islamic world. Therefore, if I had given a presentation to these leaders in the aftermath of 9/11 and incorrectly assumed that they understood the links among various militant Islamist groups, I would have completely missed the boat. Similarly, when faced with the exceedingly complicated causes of banking failure, if I write a report and assume that my audience is fully cognizant of the issues, I may fail in my presentation.

On the other hand, I may have an audience that is quite knowledgeable but not technically savvy; they do not have deep knowledge of statistical tools. Suppose that when I present them with my findings I emphasize the mathematics of my analysis; I will quickly lose my audience. Indeed, there is a difference when an expert presents to a conference of other experts as opposed to a congressional panel.

Finally, as I have stressed throughout the book, you must understand the political feasibility of your proposals. In 2009, *New York Times* columnist Thomas Friedman argued that while the Obama administration had been forthright in forcing the failing U.S. automobile giants to produce more gas-efficient vehicles, it had been absolutely "gutless" in giving American consumers incentive to buy these cars. He suggested that there should be a \$2 gasoline tax, which he called the "carbon" tax, so that the higher gas prices would induce consumers to go for the pricier fuel-efficient models. Even though Friedman's proposal is backed by impeccable logic, could it ever make much of an impression on lawmakers? No elected official, Republican or Democrat, can be expected to vote for such a measure, because raising taxes on necessity items is deeply unpopular under any circumstances. While the country was in a deep recession, what elected official would even think of proposing a doubling of the tax rate? If you, as an analyst, propose something that will go squarely against the interest of the decision makers, regardless of the cogency of your argument, you will fail to sell the results of your analysis to them.

Know What You Don't Know

The philosopher Socrates said the wise man knows what he does not know. There is nothing that can trip up an effective presentation more profoundly than the revelation

of the presenter's ignorance of any key issue. For instance, facing a rapidly escalating banking crisis in 2008, the George W. Bush administration pushed through a \$700 billion rescue plan for the failing institutions. As Treasury secretary Henry Paulson appeared before Congress to testify about the Troubled Asset Relief Program, he was asked to show where the vast amount of public money was being spent. Unfortunately, Paulson had to admit that he did not know exactly where the \$700 billion went among the banks. This gave the nation the impression that the Treasury was writing blank checks to faltering financial institutions without any accountability. On *The Daily Show*, comedian Jon Stewart mockingly asked if the folks at the Treasury knew how to use Microsoft Excel. Regardless of the cogency of Paulson's banking bailout plan, he had lost his case to the American people.

Say It with Numbers

If you want to persuade an audience with analytical knowledge, a presentation with numbers is often far more effective than one without. Numbers seem to have a magical quality about them for some people; once uttered, they take on an aura of rock-solid objectivity. During a cocktail party discussion of gay rights, someone solemnly declared that 10 percent of American men are homosexual. Immediately the discussion was altered as participants assumed the veracity of this claim. It is interesting to note that this figure came from Alfred Kinsey's pathbreaking study in 1948, when the methodologies of gathering sampling data were less scientific than they are today.⁷ Therefore, there could actually be more or fewer homosexual men in the United States, especially because changing social mores have allowed many to come out of the closet and admit their sexual orientations. Yet both sides of the debate tend to cling to the figure as irrefutable truth for political reasons. Numbers have the appearance of being precise, objective, and irrefutable. Therefore, it is no surprise that *Star Trek's* Mr. Spock, the embodiment of (at least half) human rationality, spouts the odds of survival in a seemingly impossible situation to five-digit precision. Even in our most earthly conversations, if we can phrase our arguments in terms of numbers, our chances of winning increase significantly. A most fascinating demonstration of the power of numbers at the highest level of public policy discussion can be found in the autobiography of George Ball, the undersecretary of state in the Kennedy and Johnson administrations. He writes about his experience in White House cabinet meetings as follows:

In any group where Robert McNamara (then Secretary of Defense) was present, he soon emerged as a dominant voice. I was impressed by his extraordinary self-confidence—based not on bluster but on detailed knowledge of objective facts. He gave the impression of knowing every detail of the Defense Department's vast operations and had concise and impressive views on any subject that arose, reinforcing his opinion with huge verbal footnotes of statistics. Since I am quite incapable of thinking in quantitative terms, I found McNamara's performance formidable and scintillating. He quoted precise figures, not mere orders of magnitudes. During the Vietnam War, if asked to appraise the chances of success for different optional projects, he would answer with apparent precision: one operation would have a 65 percent chance, another a 30 percent chance. Once I tried to tease him, suggesting that perhaps the chances were 64 percent and 29 percent, but the joke was not well taken.⁸

However impressed he was by the precision of McNamara's numbers, Ball laconically adds,

[McNamara's] mastery of that capricious behemoth, the defense establishment, was not achieved merely by a virtuosity with statistics: it required force of character. McNamara, moreover is a man of humanity and imagination, capable of strong commitment to cause.⁹

In other words, McNamara was simply "assessing" the probability figures in a most important policymaking body, which were taken as truth because of their seeming numerical precision and the authority and self-confidence with which they were spoken. Despite his apparent confidence, McNamara all but admitted later in his own writings that he had grave doubts about the possibility of winning the war in Vietnam.¹⁰ Therefore, it is important for us to understand the uses and misuses of numbers in policy debates. In chapter 9, I demonstrated how numbers can be manipulated to emphasize a point far beyond what is warranted by reality. Therefore, as policy analysts, we must be fully cognizant of the awful power and responsibility we share in shaping the welfare of a much larger community.

Tell a Compelling Story

The purpose of the last part of your effort is to convince decision makers of the value of your findings and to adopt policies accordingly. As an analyst, there is nothing more satisfying than knowing that your recommendations have formed the basis of an adopted policy. Therefore, your job may not be over until you have had a chance to present your case to your superiors. As you can tell, communication, oral or written, is an art form. However, I can give you some suggestions. In Appendix A of this book, I have presented an actual policy paper from the U.S. Government Accountability Office. You will do well to read my list of suggestions in the light of this excellent report by professional policy analysts.

The information age literally bombards us with innumerable pieces of information every day of our lives. As we see, listen to, and/or read them, very few get through to the realm of our conscious understanding. You may, for example, see a billboard while driving or listen to a lecture, yet a minute later you may recall absolutely nothing about the specific message that was directed at you. On the other hand, you may recall something you heard, saw, or read many years ago. The question is, what causes some messages to stick? Heath and Heath's research in mass communication has shown that a memorable message must be simple, concrete, and credible; have contents that are unexpected; appeal to our emotions; and contain a compelling storyline.¹¹ Since I am addressing a wide range of professionals who can be called "policy analysts," not all of these factors might be applicable to every situation. You must pick and choose the ones that apply to your own case.

A sticky message depends on simplicity of thought; you will remember a simple message but will soon forget a complex one. Therefore, you must make every effort at reporting your findings in the simplest of terms. In so doing, you must avoid writing in the passive voice. Although it once was fashionable for academics to write in the passive voice, because it allowed them to avoid making categorical statements, many now see

the active voice as the core of effective writing. It is also important to maintain simplicity by avoiding archaic words that your audience may not readily understand. Since your job is to make yourself as clear as possible, you should choose simple, everyday words, unless the content dictates that you choose more difficult ones.

Your message must not only be simple, it must be concrete in its conclusion. If you present your report as a list of findings, however strong, it will make little impression on your audience if not backed by a logically derived, well-considered policy recommendation. If you read the opinion pieces in the top newspapers, you will see that the authors not only present their cases effectively but most often propose some concrete steps to achieve what they are advocating.

A message is accepted not only for its contents but also for the credibility of the messenger. Sometimes such credibility comes from your authority or by virtue of the position you have within your organization. For instance, Theodore Roosevelt used to call the presidency the "bully pulpit"; because the president holds the highest office in the country, the country takes notice whenever the president speaks out. However, a president's credibility can be undermined by missteps and poor choices. For instance, in the aftermath of the devastation wrought by Hurricane Katrina, President George W. Bush praised Federal Emergency Management Agency (FEMA) director Michael Brown in front of television cameras, saying, "Brownie, you're doing a heck of a job." As the full extent of the tragedy unfolded and FEMA's ineptitude was revealed, the president lost a good bit of his credibility, and scorn was heaped on the FEMA director when he came before Congress.

The "unexpected" part of a memorable message comes when the analyst "connects the dots" for the listeners and clearly explains the confusing world in which they live. In the context of reporting your findings, empirical evidence showing a strong correspondence between the dependent and independent variables will certainly make a compelling case for you.

As human beings we remember messages that evoke emotions, particularly those that paint the portrait of an impending threat. As I discussed in chapter 1, fear is most often the primary motivator for collective action. Therefore, policy papers that reveal vulnerabilities gather far more attention and receive more urgent action than those that do not. In 1983, at the height of the Cold War, when the United States was particularly concerned about a surprise nuclear missile attack by the Soviet Union, President Ronald Reagan was approached by well-known physicist Edward Teller, who had a plan to build a ground-based laser system that would be able to shoot down incoming missiles.¹² With animation providing vivid conceptual pictures of laser beams zapping enemy ballistic missiles in flight, many were excited about the scheme, which its detractors quickly dubbed the Star Wars program. Despite the infeasibility of building such a system with the existing technology, the fear of missile warfare gripped the nation, and billions of dollars were spent on the powerful laser, which never became a reality.

Finally, memorable messages come with stories. Experimental studies show that when two similar messages are presented to an audience, one supported by statistics and the other by a suitable story, the latter inevitably sticks more than the former. Any good public speaker knows the power of a storyline. Presentations also become effective when they are accompanied by graphs, charts, pictorial renditions, and even pertinent video clips. Recent advances in technology have given us wonderful tools to create these aids.

WHOSE BALL IS IT ANYWAY? ZEN AND THE ART OF PUBLIC POLICY ANALYSIS

My analyst friend looked dejected. Shaking her head, she told me that she had spent a great deal of time preparing an analysis on a particular issue. Her results, based on a sophisticated mathematical model, clearly showed the "correct" course of action. Yet the politicians had rejected her recommendations and made an "obviously wrong decision." As I sympathized with my friend, I wondered if she should feel so dejected. Indeed, if you have spent a large amount of time on a project, it often is impossible to separate your emotions from the project itself. Her problem raises the question of the analyst's role in a public organization.¹³

At the beginning of this book, I invoked the image of the fictional fiddler on the roof, who plays his music on the slippery rooftop, performing an amazing balancing act. As policy analysts, we have more in common with him than with the white-coated scientists sequestered in sterile laboratories conducting dispassionate research. At the end of the day, our job is simply to identify the recommended path to policymakers. As analysts we would do well to remember that although the pundits play the game, within our democratic processes, it is the elected princes who really own the ball.

EXERCISES

1. Take a significant problem facing the world today. Suppose you have been asked to design a policy to deal with it. By making assumptions about your client and your resources (such as time and money), design the steps you will take to complete the task.
2. Examine the question of poverty in America. Write a brief policy paper with recommendations for a national antipoverty initiative.

Notes

1. Irving L. Janis, *Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascos* (Boston: Houghton Mifflin, 1972).
2. Dipak K. Gupta, "Tyranny of Data: Going beyond Theories," in *Tangled Roots: Social and Psychological Factors in the Genesis of Terrorism*, ed. Jeff Victoroff (Amsterdam, the Netherlands: IOS Press, 2006), 37–51.
3. Robert D. Behn and James W. Vaupel, *Quick Analysis for Busy Decision Makers* (New York: Basic Books, 1982).
4. Eugene Bardach, *A Practical Guide for Policy Analysis: The Eightfold Path to More Effective Problem Solving*, 2nd ed. (Washington, D.C.: CQ Press, 2005).
5. E. Thomas McClanahan, "Revenue Quest, Not Safety, Behind Red Light Cameras," *Kansas City Star*, January 3, 2009.
6. See Harinder Singh, "Relative Evaluation of Subjective and Objective Measures of Expectation Formation," *Quarterly Review of Economics and Business* 30 (1990): 64–74; and Roger Frantz, *Two Minds: Intuition and Analysis in the History of Economic Thought* (New York: Springer, 2005).
7. Alfred Kinsey, Wardell B. Pomeroy, and Clyde E. Martin, *Sexual Behavior in the Human Male* (Philadelphia: W. B. Saunders, 1948).
8. George W. Ball, *The Past Has Another Pattern: Memoirs* (New York: W.W. Norton, 1982), 173–174.

9. Ibid, 174.
10. Robert S. McNamara, James Blight, Robert Brigham, Thomas Biersteker, and Herbert Schandler, *Argument without End: In Search of Answers to the Vietnam Tragedy* (New York: Public Affairs, 1999).
11. Chip Heath and Dan Heath, *Made to Stick: Why Some Ideas Survive and Others Die* (New York: Random House, 2007).
12. William J. Broad, *Teller's War: The Top-Secret Story behind the Star Wars Deception* (New York: Simon & Schuster, 1992).
13. For a more detailed discussion of the role of an analyst within an organization, see Martha S. Feldman, *Order without Design: Information Production and Policy Making* (Stanford, Calif.: Stanford University Press, 1989).

Appendix A

Example of a Policy Analysis Report

GAO

United States Government Accountability Office

Report to Congressional Requesters

January 2007

POVERTY IN AMERICA

Economic Research
Shows Adverse
Impacts on Health
Status and Other
Social Conditions as
well as the Economic
Growth Rate





United States Government Accountability Office
Washington, DC 20548

January 24, 2007

The Honorable Charles B. Rangel
Chairman
Committee on Ways and Means
House of Representatives

Dear Mr. Chairman:

According to the Census Bureau, approximately 37 million people in the United States—nearly 13 percent of the total population—lived below the poverty line in 2005.¹ This percentage was significantly larger for particular population groups, specifically children, minorities, and those living in certain geographic areas such as inner cities. The federal government spends billions of dollars on programs to assist low-income individuals and families.² These programs included Medicaid, food stamps, Temporary Assistance for Needy Families (TANF), and the Earned Income Tax Credit (EITC), to name some of the largest. While some have taken issue with Census' official poverty measure and proposed alternative measures, it is generally recognized that poverty imposes costs on the nation as a whole, not merely in terms of programmatic outlays but also through lost productivity that can affect the overall economy.

A substantial body of economic research has looked at the effects of poverty and you asked us to discuss (1) what the economic research tells us about the relationship between poverty and adverse social conditions, such as poor health outcomes, crime, and labor force attachment; and (2) what links economic research has found between poverty and economic growth.

To answer these questions, we reviewed the economic literature by academic experts, think tanks, and government agencies, which we collected from searches of various databases, peer-reviewed economic journals, specialty journals, and books. We also provided our draft report

¹ In 2005 the poverty threshold for a family of four was \$19,971.

² Congressional Research Service, *Cash and Noncash Benefits for Persons with Limited Income: Eligibility Rules, Recipient and Expenditure Data, FY2002-FY2004* (Washington, D.C.: Mar. 27, 2006).

to four external reviewers. They are recognized experts who have conducted research and published on the topic of poverty and economic growth and whose work has recommended a variety of approaches and strategies to policymakers. We limited the scope of our work by looking at recent studies published since 1996, excluding anything older, with exceptions made for work that was considered seminal. Thus, our results are not an exhaustive or historical treatment of the topic. Our review was primarily driven by the economic literature focused on the United States either exclusively or including other developed nations; studies from other disciplines were excluded unless they were captured in either the economic study under review or its bibliography. When we refer to poverty in the report, we are using an absolute measure, not a relative one. This means that, for the most part, the studies we reviewed typically used the official poverty line published by the Census Bureau as its benchmark. A few of the studies we reviewed used relative measures such as the poorest 10 percent of the population.

Our work was conducted between October 2006 and January 2007 according to generally accepted government auditing standards. Because we did not evaluate the policies, operations, or programs of any federal agency to develop the information presented in this report, and because we are not making any recommendations, we did not seek agency comments. However, we met with agency officials from the Departments of Commerce, Health and Human Services, Justice, and Labor to obtain information on research they or others had conducted related to our work objectives.

Results in Brief

Economic research suggests that individuals living in poverty face an increased risk of adverse outcomes, such as poor health and criminal activity, both of which may lead to reduced participation in the labor market. While the mechanisms by which poverty affects health are complex, some research suggests that adverse health outcomes are due, in part, to limited access to health care as well as exposure to environmental hazards and engaging in risky behaviors. For example, some research has shown that increased availability of health insurance such as Medicaid for low-income mothers led to a decrease in infant mortality. Likewise, exposure to high levels of air pollution from living in urban areas close to highways can lead to acute health conditions. Data suggest that engaging in risky behaviors, such as tobacco and alcohol use, a comparatively sedentary life-style, and a low consumption of nutritional foods can account for some portion of the health disparities between lower and upper income groups. The economic research we reviewed also points to

links between poverty and crime. For example, one study indicated that higher levels of unemployment are associated with higher levels of property crime. The relationship between poverty and adverse outcomes for individuals is complex, in part because most variables, like health status, can be both a cause and a result of poverty. Regardless of whether poverty is a cause or an effect, the conditions associated with poverty can limit the ability of individuals to develop the skills, abilities, knowledge, and habits necessary to fully participate in the labor force.

Research shows that poverty can negatively impact economic growth by affecting the accumulation of human capital and rates of crime and social unrest. Economic theory has long suggested that human capital—that is, the education, work experience, training, and health of the workforce—is considered one of the fundamental drivers of economic growth. The conditions associated with poverty can work against this human capital development by limiting individuals' ability to remain healthy and develop skills, in turn decreasing the potential to contribute talents, ideas, and even labor to the economy. An educated labor force, for example, is better at learning, creating, and implementing new technologies. Economic theory suggests that when poverty affects a significant portion of the population, these effects can extend to the society at large and produce slower rates of growth. Although historically research has focused mainly on the extent to which economic growth alleviates poverty, some recent empirical studies have begun to demonstrate that higher rates of poverty are associated with lower rates of growth in the economy as a whole. For example, areas with higher poverty rates experience, on average, slower per capita income growth rates than low-poverty areas.

Background

Economic growth is one of the indicators by which the well-being of the nation is typically measured, although recent discussions have focused on a broader set of indicators, such as poverty. Poverty in the United States is officially measured by the Census Bureau, which calculates the number of persons or households living below an established level of income deemed minimally adequate to support them. The federal government has a long-standing history of assisting individuals and families living in poverty by providing services and income transfers through numerous and various types of programs.

Measuring the Nation's Well-Being: Economic Growth and Other Indicators

Economic growth is typically defined as the increase in the value of goods and services produced by an economy; traditionally this growth has been measured by the percentage rate of increase in a country's gross domestic product, or GDP. The growth in GDP is a key measure by which policy-makers estimate how well the economy is doing. However, it provides little information about how well individuals and households are faring.

Recently there has been a substantial amount of activity in the United States and elsewhere to develop a comprehensive set of key indicators for communities, states, and the nation that go beyond traditional economic measures.¹ Many believe that such a system would better inform individuals, groups, and institutions on the nation as a whole.² Poverty is one of these key indicators. Poverty, both narrowly and more broadly defined, is a characteristic of society that is frequently monitored and defined and measured in a number of ways.³

How Is Poverty Defined in the United States.?

The Census Bureau is responsible for establishing a poverty threshold amount each year; persons or families having income below this amount are, for statistical purposes, considered to be living in poverty.⁴ The threshold reflects estimates of the amount of money individuals and families of various sizes need to purchase goods and services deemed minimally adequate based on 1960s living standards, and is adjusted each year using the consumer price index. The poverty rate is the percentage of individuals in total or as part of various subgroups in the United States who are living on income below the threshold amounts.

Over the years, experts have debated whether or not the way in which the poverty threshold is calculated should be changed. Currently the calculation only accounts for pretax income and does not include noncash benefits and tax transfers, which, especially in recent years, have comprised larger portions of the assistance package to those who are low-

¹ GAO, *Informing Our Nation: Improving How to Understand and Assess the USA's Position and Progress*, GAO-05-1 (Washington, D.C., November 2004), p.4.

² GAO-05-01.

³ GAO-05-01, p.34.

⁴ The U.S. Department of Health and Human Services (HHS) establishes poverty guidelines that are similar to the poverty thresholds but are used by HHS and other agencies for administering programs, such as determining program eligibility.

income.⁷ For example, food stamps and the Earned Income Tax Credit could provide a combined amount of assistance worth an estimated \$5,000 for working adults with children who earn approximately \$12,000 a year.⁸ If noncash benefits were included in a calculation of the poverty threshold, the number and percentage of individuals at or below the poverty line could change. In 1995, a National Academy of Sciences (NAS) panel recommended that changes be made to the threshold to count noncash benefits, tax credits, and taxes; deduct certain expenses from income such as child care and transportation; and adjust income levels according to an area's cost of living.⁹ In response, the Census Bureau published an experimental poverty measure in 1999 using the NAS recommendations in addition to its traditional measure but, to date, Census has not changed the official measure.¹⁰

U.S. Poverty Rates

In 2005, close to 13 percent of the total U.S. population—about 37 million people—were counted as living below the poverty line, a number that essentially remained unchanged from 2004. Poverty rates differ, however, by age, gender, race, and ethnicity and other factors. For example,

- Children: In 2005, 12.3 million children, or 17.1 percent of children under the age of 18, were counted as living in poverty. Children of color were at least three times more likely to be in poverty than those who were white: 3.7 million, or 34.2 percent of, children who were African-American and 4 million, or 27.7 percent of, children who were Hispanic

⁷ Congressional Research Service, *Poverty in the United States: 2005* (Washington, D.C.: Aug. 31, 2006).

⁸ Danziger, Sheldon, "Fighting Poverty Revisited: What Did Researchers Know 40 Years Ago? What Do We know Today?" Dec. 4, 2006.

⁹ For a summary of the NAS panel recommendations see Congressional Research Service Report 95-539, *Redefining Poverty in the United States: National Academy of Science Panel Recommendations*, by Thomas R. Gabe (archived) (Washington, D.C.: 1995).

¹⁰ U.S. Census Bureau, *Poverty among Working Families: Findings from Experimental Poverty Measures* (Washington, D.C.: Sept. 2000).

lived below the poverty line compared to 4 million, or 9.5 percent of, children who were white.¹¹

- **Racial and ethnic minorities:** African-Americans and Hispanics have significantly higher rates of poverty than whites. In 2005, 24.9 percent of African-Americans (9.2 million) and 22 percent of Hispanics (9.4 million) lived in poverty, compared to 8.3 percent for whites (16.2 million).
- **Elderly:** The elderly have lower rates of poverty than other groups. For example, 10.1 percent of adults (3.6 million) aged 65 or older lived in poverty.

Poverty rates also differ depending on geographical location and for urban and nonurban areas. Poverty rates for urban areas were double those in suburbs, 17 percent compared to 9.3 percent.¹² Poverty rates in the South were the highest at 14 percent; the West had a rate of 12.6 percent, followed by the Midwest with 11.4 percent and the Northeast at 11.3 percent.¹³

The Role of the Federal Government

The U.S. government has a long history of efforts to improve the conditions of those living with severely limited resources and income. Presidents, Congress, and other policymakers have actively sought to help citizens who were poor, beginning as early as the 1850s through the more recent efforts established through welfare reform initiatives enacted in 1996.

Over the years, the policy approaches used to help low-income individuals and families have varied. For example, in the 1960s federal programs

¹¹ Beginning in March 2003, the Census Bureau allowed survey respondents to identify themselves as belonging to one or more racial groups. In prior years, respondents could select only one racial category. Consequently, poverty statistics for different racial groups for 2002 and after are not directly comparable to earlier years' data. The term "blacks and whites" refers to persons who identified with only one single racial group. The term "Hispanic" refers to individuals' ethnic, as opposed to racial, identification. Hispanics may be of any race.

¹² Congressional Research Service, *Poverty in the United States: 2005* (Washington, D.C.: Aug. 31, 2006).

¹³ Congressional Research Service, *Poverty in the United States: 2005* (Washington, D.C.: Aug. 31, 2006).

focused on increasing the education and training of those living in poverty. In the 1970s, policy reflected a more income-oriented approach with the introduction of several comprehensive federal assistance plans. More recently, welfare reform efforts have emphasized the role of individual responsibility and behaviors in areas such as family formation and work to assist people in becoming self-sufficient. Although alleviating poverty and the conditions associated with it has long been a federal priority, approaches to developing effective interventions have sometimes been controversial, as evidenced by the diversity of federal programs in existence and the ways in which they have evolved over time.

Currently, the federal government, often in partnership with the states, has created an array of programs to assist low-income individuals and families. According to a recent study by the Congressional Research Service (CRS), the federal government spent over \$400 billion on 84 programs in 2004 that provided cash and noncash benefits to individuals and families with limited income. These programs cover a broad array of services: Examples include income supports or transfers such as the Earned Income Tax Credit and TANF; work supports such as subsidized child care and job training; health supports and insurance through programs like the State Children's Health Insurance Program (SCHIP) and Medicaid; and other social services such as food, housing, and utility assistance.

Economic Research Links Poverty with Adverse Outcomes for Individuals Such as Poor Health and Crime

Economic research suggests that individuals living in poverty face an increased risk for adverse outcomes, such as poor health, criminal activity, and low participation in the workforce. The adverse outcomes that are associated with poverty tend to limit the development of skills and abilities individuals need to contribute productively to the economy through work, and this in turn, results in low incomes. The relationship between poverty and outcomes for individuals is complex, in part because most variables, like health status, can be both a cause and a result of poverty. The direction of the causality can have important policy implications. To the extent that poor health causes poverty, and not the other way around, then alleviating poverty may not improve health.

Individuals Living in Poverty Experience Higher Rates of Adverse Health Outcomes, in Part because of Limited Access to Health Care, Environmental Hazards, and Risky Behaviors

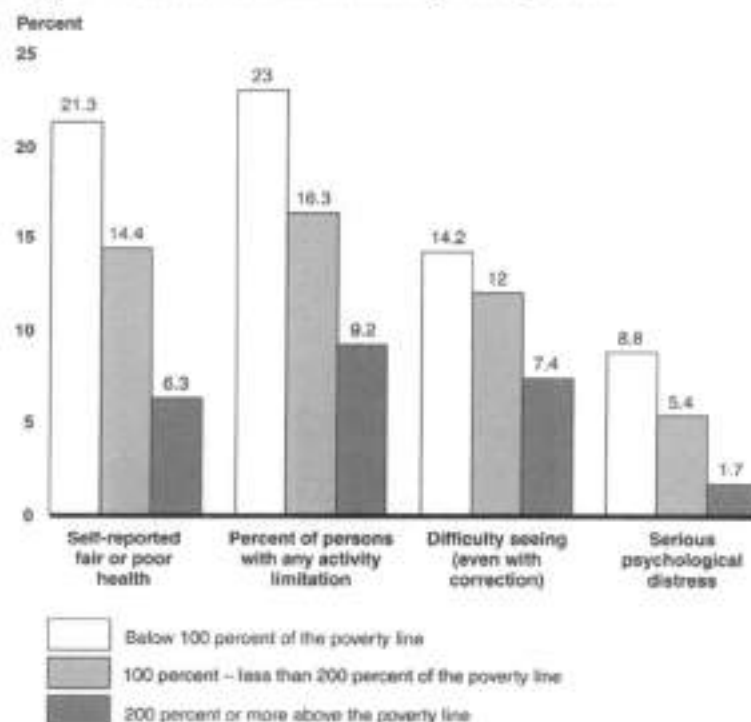
Health outcomes are worse for individuals with low incomes than for their more affluent counterparts. Lower-income individuals experience higher rates of chronic illness, disease, and disabilities, and also die younger than those who have higher incomes.¹⁴ As reported by the National Center on Health Statistics, individuals living in poverty are more likely than their affluent counterparts to experience fair or poor health, or suffer from conditions that limit their everyday activities (fig.1). They also report higher rates of chronic conditions such as hypertension, high blood pressure, and elevated serum cholesterol, which can be predictors of more acute conditions in the future. Life expectancies for individuals in poor families as compared to nonpoor families also differ significantly. One study showed that individuals with low incomes had life expectancies 25 percent lower than those with higher incomes.¹⁵ Other research suggests that an individual's household wealth predicts the amount of functionality of that individual in retirement.¹⁶

¹⁴ Centers for Disease Control and Prevention, *Health, United States, 2006; 1998* (Hyattsville, Maryland).

¹⁵ Deaton, Angus, "Policy Implications of the Gradient of Health and Wealth," *Health Affairs*, Vol. 21, No.2, March 2002.

¹⁶ Smith, James, and Raynard Kington, "Demographic and Economic Correlates of Health in Old Age," *Demography*, Vol. 34, No. 1, 1997.

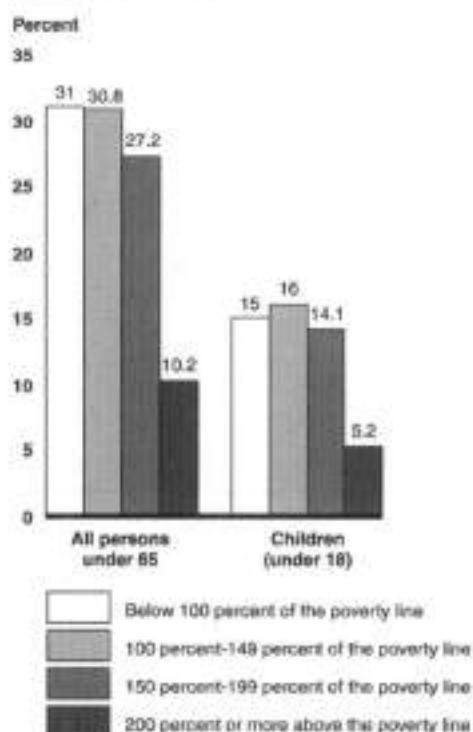
Figure 1: Selected Health Indicators by Poverty Status



Source: National Center for Health Statistics, *Health, United States, 2006 with Chartbook on Trends in the Health of Americans* (Hyattsville, Maryland: 2006).

Research suggests that part of the reason that those in poverty have poor health outcomes is that they have less access to health insurance and thus less access to health care, particularly preventive care, than others who are nonpoor. Very low-income individuals were three times as likely not to have health insurance than those with higher incomes, which may lead to reduced access to and utilization of health care (fig. 2).

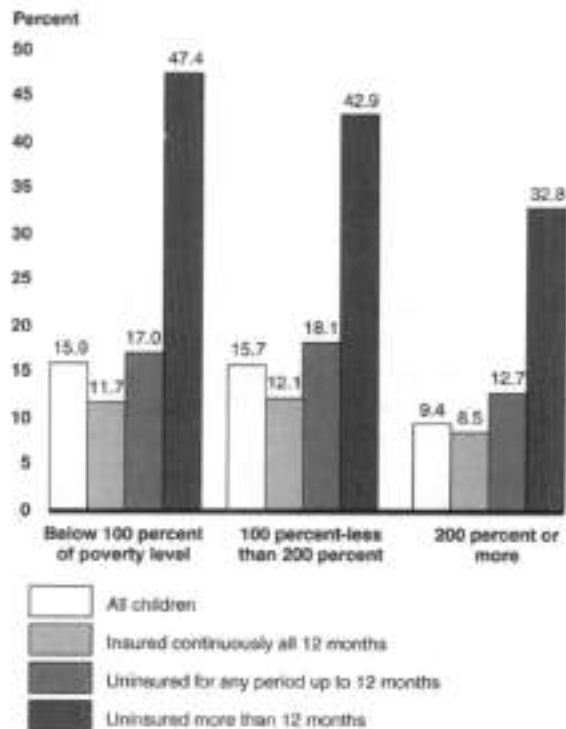
Figure 2: Percentage of Population with No Health Insurance (Private or Medicaid) by Poverty Status



Source: National Center for Health Statistics, *Health, United States, 2006 with Chartbook on Trends in the Health of Americans* (Hyattsville, Maryland: 2006).

Data show that those who are poor with no health insurance access the health system less often than those who are either insured or wealthier when measured by one indicator of health care access: visits to the doctor. For example, data from the National Center on Health Statistics show that children in families with income below the poverty line who were continuously without health insurance were three to four times more likely to have not visited a doctor in the last 12 months than children in similar economic circumstances who were insured (fig. 3). Research also suggests that a link between income and health exists independent of health insurance coverage. Figure 3 also shows that while children who are uninsured but in wealthier families visit the doctor fewer times than those who are insured, they still go more often than children who are uninsured but living in poverty.

Figure 3: No Visits to Any Health Provider in the Past 12 Months (Children under 18 Years of Age) by Level of Insurance



Source: National Center for Health Statistics, *Health, United States, 2006 with Chartbook on Trends in the Health of Americans* (Hyattsville, Maryland: 2006).

Some research examining government health insurance suggests that increased health insurance availability improves health outcomes. Economists have studied the expansion of Medicaid, which provides health insurance to those with low income. They found that Medicaid's expansion of coverage, which occurred between 1979 and 1992, increased the availability of insurance and improved children's health outcomes. For example, one study found that a 30 percentage point increase in eligibility for mothers aged 15-44 translated into a decrease in infant mortality of 8.5 percent.¹⁷ Another study looked at the impact of health insurance coverage through Medicare and its effects on the health of the elderly and also

¹⁷ Currie, Janet, and Jonathan Gruber, "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women," *The Journal of Political Economy*, Vol. 104, No. 6, December 1996.

found a statistically significant though modest impact.¹⁸ There is some evidence that variations in health insurance coverage do not explain all the differences in health outcomes. A study done in Canada found improvements in children's health with increases in income, even though Canada offers universal health insurance coverage for hospital services, indicating that health insurance is only part of the story.¹⁹

Although there is a connection among poverty, having health insurance, and health outcomes, having health insurance is often associated with other attributes of an individual, thus making it difficult to isolate the direct effect of health insurance alone. Most individuals in the United States are either self-insured or insured through their employer. If those who are uninsured have lower levels of education, as do individuals with low income, differences in health between the insured and uninsured might be due to level or quality of education, and not necessarily insurance.²⁰

Another reason that individuals living in poverty may have more negative health outcomes is because they live and work in areas that expose them to environmental hazards such as pollution or substandard housing. Some researchers have found that because poorer neighborhoods may be located closer to industrial areas or highways than more affluent neighborhoods, there tend to be higher levels of pollution in lower-income

¹⁸ Card, David, et. al., "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization and Health: Evidence from Medicare" National Bureau of Economic Research, Working Paper 10365, NBER, March 2004.

¹⁹ Currie, Janet, and Mark Stabile, "Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children," *American Economic Review*, Vol. 93, No. 5, December 2003.

²⁰ Additionally, differences in individual health outcomes can sometimes be explained by other factors that may be associated with poverty, but are difficult to detect, such as risk aversion.

neighborhoods.²¹ The Institute of Medicine concluded that minority and low-income communities had disproportionately higher exposure to environmental hazards than the general population, and because of their impoverished conditions were less able to effectively change these conditions.²²

The link between poverty and health outcomes may also be explained by lifestyle issues associated with poverty. Sedentary life-style; the use of alcohol and drugs; as well as lower consumption of fiber, fresh fruits, and vegetables are some of the behaviors that have been associated with lower socioeconomic status.²³ Cigarette smoking is also more common among adults who live below the poverty line than among those above it, about 30 percent compared to 21 percent.²⁴ Similarly, problems with being overweight and obese are common among those with low family incomes, although most prevalent in women: Women with incomes below 130 percent of the poverty line were 50 percent more likely to be obese than those with incomes above this amount.²⁵ Figure 4 shows that people living

²¹ While much of the specific biological mechanism by which air pollution might affect health is still unknown, some recent research by economists has noted a link between pollution and health, especially for infants. Currie and Neidell (2005) find that the decrease in the level of carbon monoxide in California in the 1990s had a significant effect on reducing infant mortality. See Currie, Janet, and Matthew Neidell, "Air Pollution and Infant Health: What Can We Learn From California's Recent Experience?" *Quarterly Journal of Economics*, 120 (3), 2005. Similarly, Chay and Greenstone (2003) find that the reduction in total suspended particulates due to the 1970 Clean Air Act had a significant impact on infant mortality. See Chay, Kenneth, and Michael Greenstone, "Air Quality, Infant Mortality, and the Clean Air Act of 1970," National Bureau of Economic Research, Working Paper No. 10053. NBER, 2003.

²² Institute of Medicine, Committee on Environmental Justice, "Toward Environmental Justice: Research, Education, and Health Policy Needs", (Washington, D.C.: 1999), p.6.

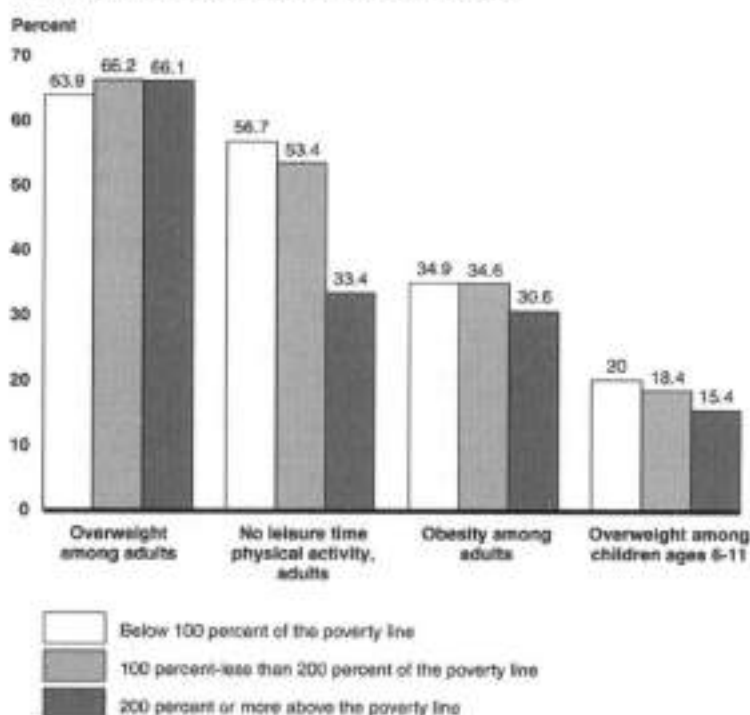
²³ Adler, Nancy E., and Katherine Newman, "Socioeconomic Disparities in Health: Pathways and Policies." *Health Affairs*, Vol. 21 No. 2, 2002. See also Deaton, Angus. "Policy Implications of the Gradient of Health and Wealth." *Health Affairs*, Vol. 21, No.2: 2002.

²⁴ Centers for Disease Control and Prevention, *Tobacco Use among Adults—United States, 2005, Morbidity and Mortality Weekly Report*, 2006; 55(42): 1145-1148. Some research suggests that part of the reason why smoking rates are higher may be peer effects, especially among youth smokers. See DeCicca, Phillip, Donald Kenkel, and Alan Mathios, "Racial Difference in the Determinants of Smoking Onset." *Journal of Risk and Uncertainty*. Boston: 2000. Vol. 21, Iss. 2/3; p311. Other studies have shown that educational attainment can affect smoking use as well. See DeCicca, Philip, Donald Kenkel, and Alan Mathios, "Putting Out the Fires: Will Higher Taxes Reduce the Onset of Youth Smoking?" *Journal of Political Economy*. Chicago 2002. Vol.110, Iss.1; p. 144.

²⁵ U.S. Public Health Service, *Surgeon General's Call To Action to Prevent and Decrease Overweight and Obesity 2001*, Washington, DC, pp. 13-14.

in poverty are less likely to engage in regular, leisure-time physical activity than others and are somewhat more likely to be obese, and children in poverty are somewhat more likely to be overweight than children living above the poverty line. In addition, there is also evidence to suggest a link among poverty, stress, and adverse health outcomes, such as compromised immune systems.²⁵

Figure 4: Percentage of Population Who Have a Sedentary Lifestyle, Are Overweight, or Are Obese, by Poverty Status



Source: National Center for Health Statistics, *Health, United States, 2009 with Chartbook on Trends in the Health of Americans* (Hyattsville, Maryland: 2009).

²⁵ While access to care, behavior, and environmental factors are some of the most commonly offered reasons for the relationship between poverty and health, recent literature has suggested other alternative theories, of which there is less of a research tradition. These include the effect of short exposures to health shocks as a result of poverty, such as poor nutrition or increased adrenalin due to higher levels of stress, and psycho-social stress that leads to problems with the immune system. See Smith, James P., "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status," *The Journal of Economic Perspectives*, Vol. 13, No. 2, 1999.

While evidence shows how poverty could result in poor health, the opposite could also be true. For example, a health condition could result, over time, in restricting an individual's employment, resulting in lower income. Additionally, the relationship between poverty and health outcomes could also vary by demographic group.²⁷ Failing health, for example, can be more directly associated with household income for middle-aged and older individuals than with children, since adults are typically the ones who work.

Economic Research Shows an Association Between Poverty and Crime

Just as research has established a link between poverty and adverse health outcomes, evidence suggests a link between poverty and crime. Economic theory predicts that low wages or unemployment makes crime more attractive, even with the risks of arrest and incarceration, because of lower returns to an individual through legal activities.²⁸ While more mixed, empirical research provides support for this. For example, one study shows that higher levels of unemployment are associated with higher levels of property crime, but is less conclusive in predicting violent crime.²⁹ Another study has shown that both wages and unemployment affect crime, but that wages play a larger role.³⁰

Research has found that peer influence and neighborhood effects may also lead to increased criminal behavior by residents. Having many peers that engage in negative behavior may reduce social stigma surrounding that behavior.³¹ In addition, increased crime in an area may decrease the

²⁷ It is not clear whether these adverse outcomes occur with greater frequency among all individuals living in households below the poverty line or only among those experiencing extreme poverty; those who experience poverty during critical development stages, such as infancy or early childhood; or those who experience long bouts of poverty.

²⁸ Criminal behavior has been measured by reports to the police in an area, self-reported crime by individuals in surveys or arrests, as well as other measures. See also Freeman, Richard, "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?" *Journal of Economic Perspectives*, Vol. 10, No. 1: Winter 2006.

²⁹ Raphael, Steven, and Rudolf Winter-Ebner, "Identifying the Effect of Unemployment on Crime," *Journal of Law and Economics*, Vol. XLIV, 2001.

³⁰ Gould, Eric D., Bruce A. Weinberg, and David B. Mustard, "Crime Rates and Local Labor Market Opportunities in the United States: 1979-1997," *Review of Economics and Statistics*, 84 (1): 2002.

³¹ Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, May 2001.

chances that any particular criminal activity will result in an arrest. Other research suggests that the neighborhood itself, independent of the characteristics of the individuals who live in it, affects criminal behavior.³² One study found that arrest rates were lower among young people from low-income families who were given a voucher to live in a low-poverty neighborhood, as opposed to their peers who stayed in high-poverty neighborhoods. The most notable decrease was in arrests for violent crimes; the results for property crimes, however, were mixed, with arrest rates increasing for males and decreasing for females.³³

Adverse Outcomes, Such as Poor Health and Low Educational Attainment, Lead to Reduced Participation in the Labor Market

Regardless of whether poverty is a cause or an effect, the conditions associated with poverty limit the ability of low-income individuals to develop the skills, abilities, knowledge, and habits necessary to fully participate in the labor force, which in turn leads to lower incomes. According to 2000 Census data, people aged 20-64 with incomes above the poverty line in 1999 were almost twice as likely to be employed as compared to those with incomes below it.³⁴ Some of the reasons for these outcomes include educational attainment and health status.

³² However, a challenge that researchers face is that, almost by definition, many individuals share the same characteristics in a neighborhood. Therefore, it is difficult to determine whether it is the characteristic of the individual or the neighborhood that is the source of the behavior.

³³ <http://www.huduser.org/publications/fairhsg/MTODemData.html> and <http://www.hud.gov/prodesc/mto.cfm>. Some economists have used data from the Moving-to-Opportunity experiment as a way to attribute causality. Moving-to-Opportunity is a research demonstration in which a number of families, chosen randomly, within five public housing authorities were given housing vouchers to be used in low-poverty neighborhoods. Another group of families acted as the control, and were not given the vouchers. Using these data, some economists have compared the outcomes for children whose families received the vouchers and those that did not. To some extent, the results have confirmed that neighborhood, independent of individual characteristics, affects criminal behavior, but the results have also been mixed. Using data from the randomized housing experiment, Ludwig, Duncan, and Hirschfeld (2001) found that the housing vouchers reduced violent arrests by teens, but may have increased the number of property arrests. Kling, Ludwig, and Katz (2005) also used the Moving-to-Opportunity data, but looked for differential effects by gender. The authors found that for females, there were large reductions in the amount of arrests for both property and violent crime, when compared to those for the control group. For males, there were reductions in violent arrests, but proportionally smaller than the drops for females. In addition, there were significant increases in the rate of property arrests.

³⁴ U.S. Census Bureau, *Employment Status: 2000, Census 2000 Brief* (Washington, D.C., August 2003), p.4

Poverty is associated with lower educational quality and attainment, both of which can affect labor market outcomes. Research has consistently demonstrated that the quality and level of education attained by lower-income children is substantially below those for children from middle- or upper-income families. Moreover, high school dropout rates in 2004 were four times higher for students from low-income families than those in high-income families.³⁶ Those with less than a high school degree have unemployment rates almost three times greater than those with a college degree, 7.6 percent compared to 2.6 percent in 2005. And the percentage of low-income students who attend college immediately after high school is significantly lower than for their wealthier counterparts: 49 percent compared to 78 percent.³⁷

A significant body of economic research directly links adverse health outcomes, which are also associated with low incomes, with the quality and quantity of labor that the individual is able to offer to the workforce. Many studies that have examined the relationship among individual adult health and wages, labor force participation, and job choice have documented positive empirical relationships among health and wages, earnings, and hours of work.³⁸ Although there is no consensus about the exact magnitude of the effects, the empirical literature suggests that poor health reduces the capacity to work and has substantive effects on wages, labor force participation, and job choice, meaning that poor health is associated with low income.

³⁶ National Center for Education Statistics, U.S. Department of Education, *Dropout Rates in the United States: 2004* (Washington, D.C. November 2006), p. 4.

³⁷ Choy, Susan, "College Access and Affordability," *Education Statistics Quarterly*, Vol. 1, Issue 2, Topic: Postsecondary Education.

³⁸ Several methodological challenges exist in this literature: For example, many of these findings could reflect the effect of income on health rather than vice versa. In addition, results are highly sensitive to the measures of health that are used, with self-reported health status subject to several forms of bias, some of which could overstate the relationship between income and health, and others of which could understate the relationship. For example, individuals who have reduced their hours of work or left the labor force may be more likely to report poor health, in order to justify their reduced labor supply or because government programs provide incentives to report disability; this would lead to an upward bias in the estimated relationship between income and health. On the other hand, it is possible that higher-income individuals, who on average have greater health care utilization, may be more likely to be diagnosed with certain conditions simply because of their greater access to health care. This would lead to a downward bias in the estimated relationship between income and health.

Research also demonstrates that poor childhood health has substantial effects on children's future outcomes as adults. Some research, for example, shows that low birth weight is correlated with a low health status later in life. Research also suggests that poor childhood health is associated with reduced educational attainment and reduced cognitive development. Reduced educational attainment may in turn have a causal effect not only on future wages as discussed above but also on adult health if the more educated are better able to process health information or make more informed choices about their health care or if education makes people more "future oriented" by helping them think about the consequences of their choices. In addition, some research shows that poor childhood health is predictive of poor adult health and poor adult economic status in middle age, even after controlling for educational attainment.

Economic Research Suggests a Negative Association between Poverty and Economic Growth

The economic literature suggests that poverty not only affects individuals but can also create larger challenges for economic growth. Traditionally, research has focused on the importance of economic growth for generating rising living standards and alleviating poverty, but more recently it has examined the reverse, the impact of poverty on economic growth. In the United States, poverty can impact economic growth by affecting the accumulation of human capital and rates of crime and social unrest. While the empirical research is limited, it points to the negative association between poverty and economic growth consistent with the theoretical literature's conclusion that higher rates of poverty can result in lower rates of growth.

Research has shown that accumulation of human capital is one of the fundamental drivers of economic growth.²⁸ Human capital consists of the

²⁸Economic models that consider human capital to be a fundamental driver of economic growth are commonly referred to as endogenous growth models, although the more traditional neoclassical model has also been augmented to include the role of human capital. Endogenous growth theory posits technological growth as occurring through dynamics inside the model. Although there are several competing models, crucial importance in each is given to the production of new technologies and human capital. While the major point these models emphasize is that human capital is the driving force behind growth, the actual modeling of the relationship is still a controversial issue in the economic literature. Some growth models assert that the driving force behind economic growth is the rate of accumulation of human capital, in which the rate of economic growth is proportional to the rate of accumulation of human capital. Another approach considers that high levels of human capital, as embodied in the level of the educational attainment of the workforce, increases the capacity of individuals to innovate (discover new technology) or to adopt new technology.

skills, abilities, talents, and knowledge of individuals as used in employment. The accumulation of human capital is generally held to be a function of the education level, work experience, training, and healthiness of the workforce.³⁰ Therefore, schooling at the secondary and higher levels is a key component for building an educated labor force that is better at learning, creating, and implementing new technologies. Health is also an important component of human capital, as it can enhance workers' productivity by increasing their physical capacities, such as strength and endurance, as well as mental capacities, such as cognitive functioning and reasoning ability. Improved health increases workforce productivity by reducing incapacity, disability, and the number of days lost to sick leave, and increasing the opportunities to accumulate work experience. Further, good health helps improve education by increasing levels of schooling and scholastic performance.

The accumulation of human capital can be diminished when significant portions of the population have experienced long periods of poverty, or were living in poverty at a critical developmental juncture. For example, recent research has found that the distinct slowdown in some measures of human capital development is most heavily concentrated among youth from impoverished backgrounds. When individuals who have experienced poverty enter the workforce, their contributions may be restricted or minimal, while others may not enter the workforce in a significant way. Not only is the productive capability of some citizens lost, but their purchasing power and savings, which could be channeled into productive investments, is forgone as well.

In addition to the effects of poverty on human capital, some economic literature suggests that poverty can affect economic growth to the extent that it is associated with crime, violence, and social unrest. According to some theories, when citizens engage in unproductive criminal activities they deter others from making productive investments or their actions force others to divert resources toward defensive activities and expenditures. The increased risk due to insecurity can unfavorably affect

³⁰ In general, economists regard expenditures on education, training, medical care, and so on as investments in human capital. Collectively, theoretical growth models suggest economic growth results from improvements in human capital as embodied in the skills and experience of the labor force; from expansion of physical capital in the form of plant and equipment; and from progress in science, engineering, and management that generates technological advance. While many variables have been empirically tested, only a few have been accepted as being statistically significant in explaining growth. The role of human capital is now almost universally regarded as being indispensable in this respect.

investment decisions—and hence economic growth—in areas afflicted by concentrated poverty. Although such theories link poverty to human capital deficiencies and criminal activity, the magnitude of their impact on economic growth for an economy such as the United States is unclear at this time.⁴⁰ In addition, people living in impoverished conditions generate budgetary costs for the federal government, which spends billions of dollars on programs to assist low-income individuals and families. Alleviating these conditions would allow the federal government to redirect these resources toward other purposes.

While economic theory provides a guide to understanding how poverty might compromise economic growth, empirical researchers have not as extensively studied poverty as a determinant of growth in the United States. Empirical evidence on the United States and other rich nations is quite limited, but some recent studies support a negative association between poverty and economic growth. For example, some research finds that economic growth is slower in U.S. metropolitan areas characterized by higher rates of poverty than those with lower rates of poverty.⁴¹ Another study, using data from 21 wealthy countries, has found a similar negative relationship between poverty and economic growth.⁴²

Concluding Observations

Maintaining and enhancing economic growth is a national priority that touches on all aspects of federal decision making. As the nation moves forward in thinking about how to address the major challenges it will face

⁴⁰ Human capital deficits experienced by some impoverished individuals cannot always be attributed to experience of poverty. In some cases, low education attainment and poor health, although associated with poverty, may actually be caused by some other factor that is also responsible for poverty. In this case, poverty would be a symptom rather than a cause (i.e., poor health, poor choices, or addiction may erode human capital potential and cause poverty). Similarly, most poor people do not commit crimes, and those that do may be motivated by forces unrelated to their incomes.

⁴¹ The relationship is not always statistically significant in all regions. Statistical insignificance in some cases might be more attributable to data issues such as sample size or multicollinearity rather than an indication of nonrelationship between poverty and income growth in various regions. See S. Dev Bhatta, "Are Inequality and Poverty Harmful for Economic Growth," *Journal of Urban Affairs*, 22 (3-4): 2001. This study provides, arguably, a better comparison group than cross-country studies, since metropolitan statistical areas in the United States are at relatively similar stages of development.

⁴² Voitchovsky, S., "Does the Profile of Income Inequality Matter for Economic Growth? Distinguishing between the Effects of Inequality in Different Parts of the Income Distribution," *Journal of Economic Growth*, Vol. 10: 2005.

in the twenty-first century, the impact of specific policies on economic growth will factor into decisions on topics as far ranging as taxes, support for scientific and technical innovation, retirement and disability, health care, education and employment. To the extent that empirical research can shed light on the factors that affect economic growth, this information can guide policymakers in allocating resources, setting priorities, and planning strategically for our nation's future.

Economists have long recognized the strong association between poverty and a range of adverse outcomes for individuals, and empirical research, while limited, has also begun to help us better understand the impact of poverty on a nation's economic growth. The interrelationships between poverty and various adverse social outcomes are complex, and our understanding of these relationships can lead to vastly different conclusions regarding appropriate interventions to address each specific outcome. Furthermore, any such interventions could take years, or even a generation, to yield significant and lasting results, as the greatest impacts are likely to be seen among children. Nevertheless, whatever the underlying causes of poverty may be, economic research suggests that improvements in the health, neighborhoods, education, and skills of those living in poverty could have impacts far beyond individuals and families, potentially improving the economic well-being of the nation as a whole.

Appendix B

Models of Causal Prediction: Multiple Regression

If you are interested in forecasting an event, you may do so by extending the trend or by examining the event's root causes. Forecasting on the basis of trend was covered in chapter 12. This appendix covers causal prediction, the method to be used when models of trend projection may be inappropriate.

As you have learned, trend projection models assume that the present trend will continue, at least for the period of prediction. The problem with an assumption of continuity is that a past trend may not hold true for any length of time. Or the data may not show any obvious trend pattern. In these cases we must think about the causal relations that may link the dependent variable with a set of independent variables.

Causal models based on multiple regression techniques have been used in public policy analyses in many different ways. Many local governments depend on multiple regression models for assessing property taxes.¹ Property tax is assessed on its market value. However, we cannot know the value before a property is sold, so an assessor must estimate, or forecast, the market price.² Because a high assessment means a high property tax for the homeowner, property tax assessments have been a source of conflict in many communities across the nation. How do you forecast a property's market price? Let us look at a hypothetical example. Suppose your town's appraiser has assessed the value of a custom-built home at \$265,000, which the owner is challenging as far above its market value. Your job is to determine the validity of the owner's claim. You will have to build a causal model for predicting the market value of the property.

BUILDING A CAUSAL MODEL

The first step in building a causal model requires establishing the causal linkage between a dependent variable and a set of independent variables. Once we have inquired into the causes of a specific event (the dependent variable), we begin our analysis by selecting the independent variables. After preparing a list and collecting data, we determine how these independent variables relate to the dependent variable. These two steps constitute the specification of the model. At this stage, the difference between a time-series model and a causal model comes into sharper focus. For building a model in which the only explanatory factor is the trend over time, we do not need to understand the causes of a

change. However, in specifying a model for causal analysis, we must develop a deep understanding of the change's causes.

Let us return to our example of predicting the price of a house. As the first step toward specification, we hypothesize that the price of a house depends on the following independent variables: the type of structure and its location. We assume that the type of structure has only one component: the size of the covered area. The location factors include two separate variables: whether the house has a view and whether it is located on a cul-de-sac. You can write your model as

$$\text{Price} = f(\text{Area, View, Cul-de-sac}) \quad (\text{B.1})$$

Expression B.1 is read as "price is a function of (or depends on) area, view, and cul-de-sac." This expression is also known as an implicit function or implicit model. It is called implicit because we are not making any explicit hypotheses regarding the nature of the relationship between the dependent variable and the independent variables. If we express the model by specifying the relationship, the expression is called an explicit function or explicit model. Thus, the implicit function written in equation B.1 can be made explicit by writing

$$\beta_0 + \beta_1 \text{Area} + \beta_2 \text{View} + \beta_3 \text{Cul-de-sac} + \varepsilon \quad (\text{B.2})$$

where β_i are the coefficients for the independent variables, and ε is the error term.

The explicit model (equation B.2) states that we expect the price to go up with the size and relative desirability of the area and location factors. Also, the model states that we are hypothesizing a linear relationship between price and the other independent variables.

CAUSALITY VERSUS CO-OCCURRENCE

One of the most difficult problems of statistical analysis based on regression or a correlation coefficient is that the results can never establish causality; they can only demonstrate co-occurrence. That is, empirical results can establish that the dependent and the independent variables moved in the same direction, but they cannot say whether one caused the other. Actual causality must be established by a theoretical explanation of human behavior. Policy researchers develop their hypotheses by thoroughly reviewing the literature on the theories of human social interactions.

An example may clarify this point. Data show that productivity in the United States, measured in per capita gross domestic product (GDP) and adjusted for inflation (real per capita GDP), is increasing steadily over the years. So is the population of Lima, Peru. If we regress U.S. per capita GDP against the population of Lima, we will find a high degree of statistical significance, leading to the claim that an increase in Lima's population adds to U.S. productivity. Yet it is obvious that no such causal connection exists. On the other hand, if you study the literature on economic growth, you find that the accepted theories explain growth in national per capita GDP by pointing to a combination of expenditure on financial capital (total national capital formation) and spending on human capital (for example, education, health care, or research and development).

The issues of causality and correlation are central to our understanding of the methods of scientific research. Hardly a day passes without medical researchers' linking some food or personal habit with some human ailment. Many of these findings, of course, cannot claim causality and instead point out close associations. Sometimes researchers misuse these findings by asserting a causal connection where there is none. In contrast, some people take advantage of the fact that these results do not *prove* causality and deny a significant relationship. Lobbyists for the tobacco industry have tried to dismiss the results linking lung cancer with smoking as a mere co-occurrence. However, in the face of overwhelming evidence, increasing public awareness has prompted public policies that limit tobacco use.

Other than mistaking co-occurrence for causality, we may make significant mistakes by using trivial variables to find a causal linkage. Suppose we want to develop a model for explaining the difference in criminal justice expenditures among cities of similar size. If we include in the model the number of law enforcement officers as an independent variable, we will be engaging in a trivial pursuit, since it is obvious that when we hire more officers, the size of our budget goes up. Instead, we should use variables such as crime rate, population density, and per capita city income. Of course, when trying to exclude trivial independent variables, lines tend to blur. A deep understanding of human behavior, not statistical manipulation, helps you see that link between the dependent and independent variables.

ESTIMATION OF THE MODEL

After specifying a model, we estimate the relevant coefficients. For this step we need to determine how each of the variables is to be quantified or operationalized. Thus, in our example, we can operationalize housing quality as an explanation of price with the help of our independent variables: the size of a house expressed in square feet of covered area, the presence or absence of a view, and the location on or off a cul-de-sac (which is desirable for privacy and the absence of fast-moving traffic). We expressed each location variable (view and cul-de-sac) as an intercept dummy, with a value of 1 if a particular house has a view and 1 if it is located on a cul-de-sac. If, on the other hand, it has no view and is not on a cul-de-sac, each location variable gets a value of 0.

Suppose the assessor's office has provided you with information on thirty houses in the neighborhood sold during the past six months. The information is presented in Table B.1. You can estimate the model with these data.³

Using information from Table B.1, you can write the estimated equation as follows:

$$\begin{aligned}
 \text{Price} &= 6,767.87 + 100.04 \text{ Area} + 40,603.07 \text{ View} + 42,607.31 \text{ Cul-de-sac} \\
 t &= (0.233) \quad (8.18)^* \quad (3.23)^* \quad (3.38)^* \\
 R^2 &= 0.786 \\
 \text{Adjusted } R^2 &= 0.761 \\
 F(3, 26) &= 31.77^* \\
 n &= 30
 \end{aligned}
 \tag{B.3}$$

*Significant at the .05 level

TABLE B.1 Housing Prices and Characteristics

	Price (\$)	Area (square feet)	View	Cul-de-sac
1	310,000	2,200	1	1
2	233,000	1,800	0	0
3	400,000	3,500	1	0
4	430,000	3,200	1	1
5	210,000	1,800	0	0
6	240,000	1,700	1	1
7	300,000	2,200	0	1
8	350,000	2,100	1	1
9	385,000	2,600	1	1
10	368,000	3,000	0	0
11	200,000	2,000	0	0
12	298,000	1,750	1	1
13	275,000	1,900	0	1
14	198,000	1,800	0	0
15	253,000	2,200	1	0
16	278,000	2,100	1	1
17	320,000	2,170	1	1
18	178,000	1,200	1	1
19	225,000	2,000	0	0
20	212,000	1,900	0	1
21	288,000	2,800	0	1
22	315,000	2,300	1	1
23	255,000	2,600	0	1
24	284,000	2,700	1	0
25	189,000	1,640	1	0
26	220,000	2,600	0	0
27	248,000	1,900	1	0
28	276,000	2,000	1	1
29	210,000	2,300	0	0
30	205,000	2,200	0	0

The estimated equation (B.3) tells an interesting and convincing story about how home prices are determined. It states that each square foot adds \$100.04 to the price of a home. The estimated results also point out that location variables contribute heavily to the determination of price. If the property is located on a piece of land with a view, the price increases by \$40,603.07, while sitting on a cul-de-sac boosts the price by another \$42,607.31.

The interpretation of the intercept term is often problematic in a regression equation. Equation B.3 indicates that even if there is not a single square foot of covered area, there is still a price. Therefore, the temptation to interpret the intercept term as the price of land is great. Using this logic, we can state that if a parcel of vacant land (square feet = 0) does not have a view (view = 0) and is not located on a cul-de-sac (cul-de-sac = 0), it will still carry a price of \$6,767.87. With a view, this price will change to \$47,370.94. If the land is located on a cul-de-sac but has no view, it is worth \$49,375.18. Finally, when a plot of land in this particular neighborhood has a view and

is located on a cul-de-sac, it should command a price of \$89,978.25. The danger in such an assertion is that we can interpret the intercept term to be the land price if we accept the model to be perfectly specified. If not, we may have left out one or more significant independent variables, the impact of which will then be captured by the intercept term. For example, we did not use the lot size (or the total land area) for the houses. Since this is a potentially important variable, we cannot readily interpret the intercept term. Therefore, unless the model is fully specified (includes all important variables), any interpretation of the intercept becomes questionable.

The estimated results indicate that the four independent variables explain 78.6 percent of the variations (R^2). Since there are 30 observations, at 29 degrees of freedom, the relevant t value for evaluating the significance of the four estimated coefficients (including the constant) at a 95 percent level of confidence is 1.697. All coefficients except for the constant have t values greater than this number, so all are statistically significant. The critical F value at 3 and 26 degrees of freedom is 2.76. Since the calculated F value is 31.768, we can safely conclude that the equation as a whole is statistically significant.

Now that you are satisfied with the quality of the estimated model, you can predict the price of the house in dispute. The equation states that each square foot of housing area adds \$100.04 to the base price of \$6,767.87 (value of the intercept). Also, if a house has a view, its price goes up by \$40,603.07. Finally, its location on a cul-de-sac increases its value by \$42,607.31. This particular house has 2,200 square feet of covered area and, although it does not have a view, it is located on a secluded cul-de-sac. Armed with the information derived from the estimated model, you can forecast the market price as follows:

$$\text{Estimated price} = 6,767.87 + (100.04 \times 2,200) + (40,603.07 \times 0) + (42,607.31 \times 1)$$

Notice that since this house does not have a view, the value of the dummy variable is 0. But because the house is located on a cul-de-sac, the dummy value of this variable is 1. Using this equation, we calculate the estimated price of the property to be approximately \$269,463. Therefore, according to your estimation, the owner does not have a strong case. In fact, from the estimated result, we learn that the property has been appraised at a value lower than its estimated market price.

HOW GOOD IS THE MODEL?

Because of its mathematical elegance and ease of explanation, multiple regression is probably the most commonly used technique in statistical research and econometric prediction. Although regression analysis offers a powerful tool for prediction, its quality depends on the model and how close it is to the "true" relationship. Again, as we have mentioned before, we cannot observe the "truth." Therefore, we have no objective way of knowing whether we are sufficiently close to a desired specification. Since the process of specification contains two separate parts, choosing the right set of independent variables and postulating the correct functional form, let us discuss both parts in detail, beginning with the problem of selecting the right set of independent variables.

Choosing Independent Variables

When choosing independent variables, we might omit important variables or include some that are irrelevant. The error of omitted variables occurs because of our lack of understanding of the underlying factors that produce the dependent variable. Such lack of understanding may result from a general dearth of theoretical insight into a particular aspect of social behavior or from a paucity of data. If we want to determine the reasons students engage in random school shootings, we may find out that there is neither a strong theoretical explanation for this puzzling behavior nor a good deal of data on potential offenders to specify a causal model. Therefore, if we attempt to build a model, we are likely to encounter the problem of misspecification and are apt to leave out important explanatory variables or include irrelevant ones. Either error leads to some rather serious problems of estimation. Let me explain. The basic idea of a multiple regression is illustrated in Figure B.1.

As you can see, I have explained the dependent variable, housing price, with the help of two independent variables, view and area. The intersecting parts of the Venn diagram show which aspects of the dependent variable are explained by the two independent variables. The regression model explains the area where the dependent and independent variables overlap. As you can see, the area of overlap is divided into three segments. The area V signifies the price difference among the houses explained by the presence or absence of a view. Segment A is the price that is explained by the size of the structure (area). However, there is also an area of correlation between the size of the house and the view. The correlation may be high, since the lots with prime views

FIGURE B.1 Visual Explanation of Multiple Regression

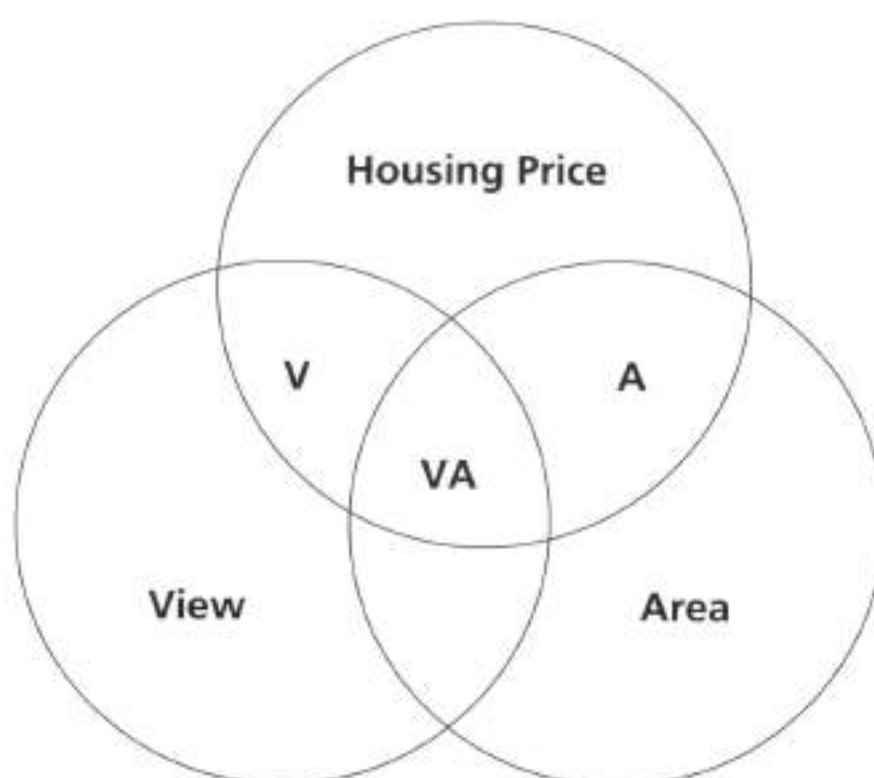
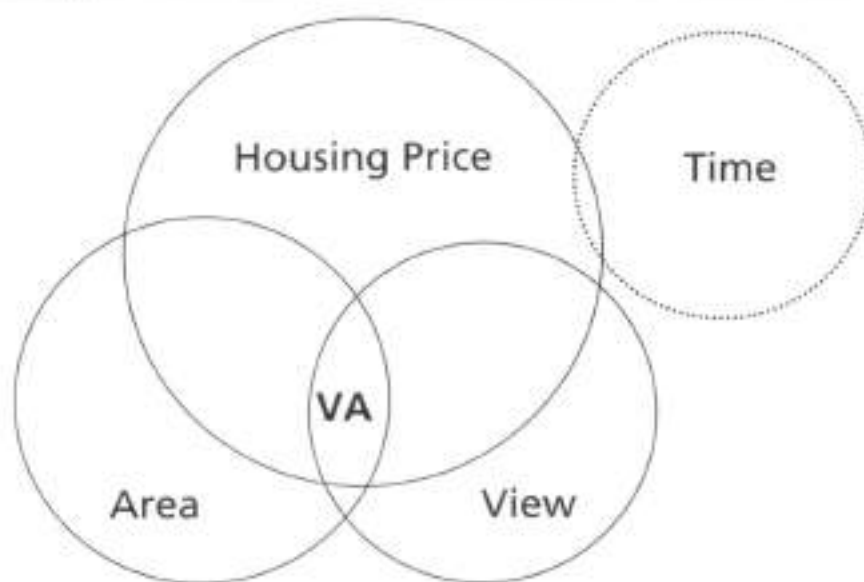


FIGURE B.2 Visual Explanation of Irrelevant Variable


tend to attract larger homes. This area is designated VA. The R^2 value derived from the estimated equation is a ratio of the total area of these three sections to the total area of the dependent variable $[(A + V + VA)/\text{housing price}]$. A high R^2 value would mean that these overlapping areas cover most of the dependent variable, and a low R^2 value would imply a small area of intersection.

We would misspecify a model if we excluded an important variable and/or included an irrelevant one. Thus, in the case of estimating housing prices, if we leave out either the area or the view, we will run into the problem of misspecification. A small R^2 value is the first indication of misspecification (the incorrect construction of a regression model) due to an omitted variable.

On the other hand, if we include irrelevant variables, the overall explanation of the dependent variable will not improve. In Figure B.2, the irrelevant variable (time) is shown with a dotted line. You may notice that compared with the other two independent variables, this one explains very little of the dependent variable (housing price). Therefore, its inclusion has little or no effect on the total explanation (R^2) of the model. Also, when we compare the adjusted R^2 value of the new model with that of the old model (without the variable time), the adjusted R^2 value of the expanded model turns out to be smaller than that of its properly specified predecessor.

Let me now explain the impact of misspecification on t values. Suppose in our housing price example that the true relation is captured by the specification of equation B.2; that is, housing prices depend on these three variables: area, view, and cul-de-sac. In this case the estimated coefficients will reflect their true values. For the purpose of exposition, let us say that we did not include the independent variable area. As a result, equation B.2 is improperly specified and is written as

$$\beta'_0 + \beta'_1 \text{View} + \beta'_2 \text{Cul-de-sac} + \varepsilon \quad (\text{B.4})$$

Since the true relationship includes area (see equation B.3), the effects of the omitted variable will be distributed among the remaining independent variables and error term. This situation will create two problems. First, the estimated coefficients for the independent variables, by illicitly incorporating fractions of another coefficient, will deviate from their true value. Second, the error term will now contain bias due to contamination by the residual factors of the omitted variable, causing the regression to violate the precepts of the classical least squares method. Moreover, the estimated coefficients will not be the best, linear, unbiased estimators. Let us see what happens to our estimated equation as a result of errors of omission. First we will omit an important explanatory variable, and then we will include an irrelevant one. We have reestimated equation B.4 by omitting the variable area from equation B.3, with the following results:

$$\begin{aligned}
 \text{Price} &= 230,560.40 + 43,030.87 \text{ View} + 34,230.87 \text{ Cul-de-sac} \\
 t &= (12.72)^* \quad (1.84)^* \quad (1.47) \\
 R^2 &= 0.234 \\
 \text{Adjusted } R^2 &= 0.177 \\
 F(2, 27) &= 4.11^* \\
 n &= 30
 \end{aligned}
 \tag{B.5}$$

*Significant at the .05 level

If you compare the estimated results of equation B.5 with those of equation B.3, you can see the obvious differences. First, both the R^2 and the adjusted R^2 values have dropped significantly, implying that less is explained by the new model. Second, quite significantly, the variable cul-de-sac is no longer statistically significant. To a researcher, these are the telltale signs of a poorly specified model. Therefore, any prediction based on such a model would be erroneous.

Errors of specification, however, are the most difficult to detect. If you have omitted some important explanatory variables, the first indication may be a lower than expected R^2 value. Because R^2 measures the percentage of explained variance, leaving out important explanatory variables will cause it to be low.

As we have seen, the omission of an important variable also causes its effect to be absorbed by the error term. The second probable sign of omission is a high serial correlation, in which the error terms of one period are correlated with those of previous periods. The problem of serial correlation will be discussed later, but it may suffice at this time to point out a few important facts. You may recall that one of the fundamental assumptions of the least squares method is the random distribution of error terms. Now suppose that the properly specified model for determining housing prices consists of the three independent variables, area, view, and cul-de-sac, but that we have failed to include cul-de-sac. In that case, the misspecified model is

$$\beta''_0 + \beta''_1 \text{Area} + \beta''_2 \text{View} + \varepsilon \tag{B.6}$$

Since I have omitted cul-de-sac, its effects are now incorporated in the error term (ε), which, in effect, has become $\varepsilon + \beta''_3 \text{cul-de-sac}$. As you can imagine, because the variable cul-de-sac is causally linked with price, the effects of its omission will not be random, and its inclusion in the error term will cause a bias. This bias in the error term causes serial correlation. If there is evidence of a high degree of serial correlation, you should look for important explanatory variables that you may have inadvertently left out.

Let me repeat: the error of omitted variables can be corrected not by resorting to any statistical technique but through an understanding of the nature of the dependent variable. This understanding must come from the theoretical literature. If we are trying to build a model to predict the crime rate of a city, we must have a deep understanding of the sociological factors that determine the overall crime rate. In forecasting revenue for the state government, we must look into economic theory and understand the state's fiscal structure.

We noted earlier that the second source of specification error is the inclusion of irrelevant variables in a model. What happens when we include independent variables that do not explain the dependent variable? This error is easier to detect than an error of omission and even simpler to correct (by eliminating irrelevant variables from the model). Let us see what happens when we include irrelevant variables in the model.

Suppose that in estimating housing prices, we decide to include a fourth independent variable, which measures the ease of access from a house. This variable, shown in Table B.2, is measured by the minutes of travel time to reach the nearest highway.

When we include this new variable and reestimate our model, we obtain the following results:

$$\begin{aligned}
 \text{Price} &= 9,291.11 + 99.78 \text{ Area} + 41,496.06 \text{ View} + 43,182.00 \text{ Cul-de-sac} - 180.5 \text{ Time} \\
 t &= (0.30) \quad (8.0)^* \quad (3.15)^* \quad (3.32)^* \quad (-0.29) \\
 R^2 &= 0.786 \\
 \text{Adjusted } R^2 &= 0.752 \\
 F(4, 25) &= 23.01^* \\
 n &= 30
 \end{aligned}
 \tag{B.7}$$

*Significant at the .05 level

As you can see from the preceding results, the new variable turns out to be statistically insignificant. This lack of statistical significance may be due to some people preferring easy access to the highway, while others prefer seclusion and distance from it. Therefore, these conflicting preferences do not show up in the determination of demand for housing. Returning to Figure B.2, you can see that when we include an irrelevant variable (time), the area of explanation (where time overlaps housing price) is rather small, and hence its inclusion does not significantly increase the R^2 value.

The inclusion of irrelevant variables poses less of a specification problem than the omission of important variables. Thus, if a new variable is totally random, it will impose no bias on either the estimated coefficients or the error term. However, unless

TABLE B.2 Variables for Determining Housing Price

Price (\$)	Area (square feet)	View	Cul-de-sac	Time (minutes)
310,000	2,200	1	1	25
233,000	1,800	0	0	3
400,000	3,500	1	0	17
430,000	3,200	1	1	8
210,000	1,800	0	0	2
240,000	1,700	1	1	35
300,000	2,200	0	1	7
350,000	2,100	1	1	33
385,000	2,600	1	1	35
368,000	3,000	0	0	21
200,000	2,000	0	0	10
298,000	1,750	1	1	1
275,000	1,900	0	1	15
198,000	1,800	0	0	4
253,000	2,200	1	0	15
278,000	2,100	1	1	13
320,000	2,170	1	1	9
178,000	1,200	1	1	32
225,000	2,000	0	0	22
212,000	1,900	0	1	25
288,000	2,800	0	1	1
315,000	2,300	1	1	7
255,000	2,600	0	1	19
284,000	2,700	1	0	20
189,000	1,640	1	0	20
220,000	2,600	0	0	4
248,000	1,900	1	0	3
276,000	2,000	1	1	15
210,000	2,300	0	0	10
205,000	2,200	0	0	25

you are deliberately choosing a random series, few data sets in nature will be randomly distributed with respect to another, especially when a researcher has reason to believe that there is *some* causal link between the dependent and the independent variables. This largely irrelevant variable will impose some biases and can be detected. To locate this variable, look for the following results:

1. If the included variable is not significantly correlated with the dependent variable, this lack of correlation will show up as an insignificant t value for the included variable's estimated coefficient. As can be seen (equation B.7), the variable time has a statistically insignificant t value.
2. Although the R^2 value will remain unchanged or may even go up slightly, the adjusted R^2 value will register a decline. You may recall our discussion of R^2 and adjusted R^2 in chapter 12. R^2 is the measure of total explanation. Therefore, even

if you add an utterly irrelevant variable, the worst it can do is to add nothing to the explanation. In this case R^2 will remain the same (unchanged) when you include such a variable in the original list of independent variables. However, you may recall that the formula for adjusted R^2 is

$$\overline{R^2} = 1 - (1 - R^2) \frac{n-1}{n-k}$$

So unless the new variable adds more to the value of R^2 than does the correction factor, $(n-1)/(n-k)$, made for the inclusion of an additional variable, the value of R^2 will decline. By comparing the estimation results of equation B.3 with those of equation B.7, we can see that the R^2 value has remained unchanged at 0.786 with the inclusion of the irrelevant variable, time. However, since this marginal increase could not compensate for the adjustment factor, the adjusted R^2 value has declined from 0.761 to 0.752.

3. The inclusion of irrelevant variables tends to increase the variance of the estimated coefficients. This increased variance reduces the t values of the coefficients. As you can see, the inclusion of an irrelevant variable has caused a slight reduction in the t values of all three independent variables.
4. An irrelevant variable may mean that the new variable is measuring the same phenomenon as one of the other independent variables. In this case, the analysis suffers from multicollinearity, a problem discussed at greater length later.

Searching for the Proper List of Independent Variables

Among the greatest beneficiaries of computer hardware and software advances have been econometricians. Thanks to the increased computing capabilities of modern computers, the most sophisticated techniques of estimation are available to prospective users at the touch of a button. As a result, the temptation to forgo the thoughtful but time-consuming effort required to build a theoretically sound model and to settle for a model of empirical convenience is strong. Yielding to it can often lead to serious but undetected problems of specification error. To underline the fact that there is no substitute for the deep understanding of the causal interrelationship, I will discuss some of the most commonly misused techniques for solving specification problems.

Suppose I am trying to build a causal model for forecasting a complex sociological phenomenon: drug abuse cases among high school students. An industrious researcher, I have found information on various social and economic characteristics of the general population in the surrounding areas of the school district. However, lacking familiarity with the scholarly literature on the subject, I do not have a clue about which variables are important and which are not. But it is easy to run a regression equation when the data are already in the machine, so there is no stopping me. I proceed to run a large number of equations and then choose the combination that gives me the highest R^2 and best t values. This approach is called a fishing expedition or data mining. The problem is that since the model has been built without a profound comprehension of the causal linkages, what we may be observing is a simple case of co-occurrence. Therefore, this model is likely to give us misleading forecasts, as the future development of these less

than relevant independent variables will have little bearing on the course of the dependent variable.

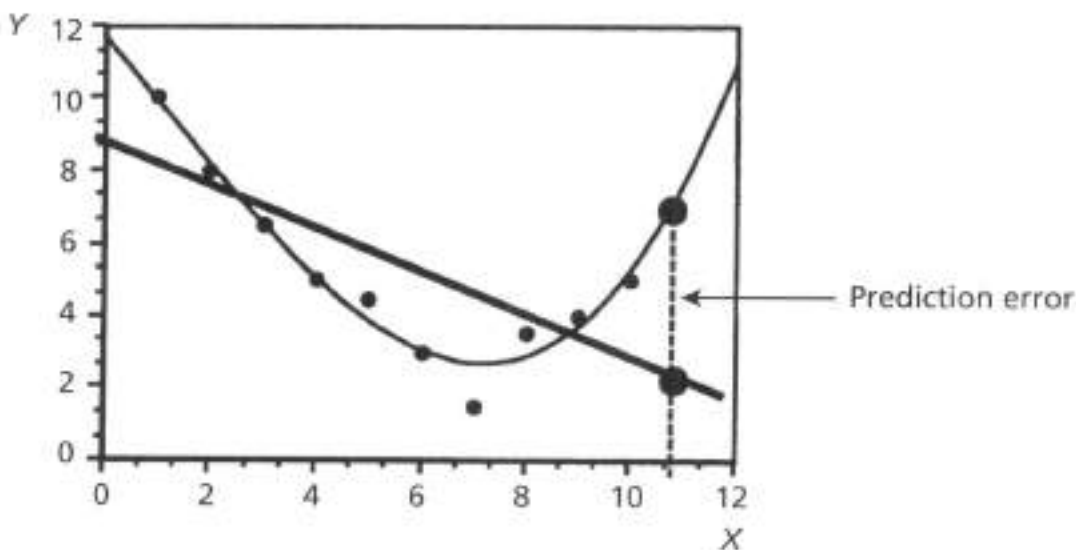
Stepwise regression is a statistical technique that minimizes the tedious job of having to choose manually the best set of independent variables. Given a list of variables, this technique will search for the variable that gives the highest value of R^2 . After that, it will pick from the list the second variable that adds the most to this R^2 value, then go to the third, and so on. Many statistical packages come with this stepwise option. The problem with stepwise regression is that it selects independent variables solely on the basis of the strength of association, not on causality.

Predicting on the Basis of the Wrong Functional Specification

Even after we have selected the “correct” set of independent variables, serious errors can arise if we do not choose the right functional form. For example, suppose we have a data set in which the true relationship is a quadratic one. If we assume that the relationship is linear and fits a straight line, we will get a terribly inaccurate predicted value (see Figure B.3). I have discussed this problem in chapter 12.

The only way to avoid the problem of misspecifying the functional form is to graph each independent variable against the dependent variable. Although a linear form is recommended over a more complicated nonlinear one, you cannot always use the linear functional form. If the data are arranged in a way that quickly reveals, say, a quadratic form, you may not have to spend time plotting the data. However, in most cases you cannot be sure until you have plotted them. Consider the two sets of data presented in Table B.3. Series A presents citywide survey results showing support for a public roller-skating

FIGURE B.3 Errors of Functional Misspecification



Note: The figure shows a huge error in prediction due to the misspecification of the functional form. Clearly, in this case, the polynomial form is more appropriate. The choice of a straight line causes prediction error.

TABLE B.3 Detecting the Relationship between Variables in a Series

Community support for roller-skating park

Series A (1991–2000)

Series B (2001)

Year	Community support (percent)	Household median income ($\times \$1,000$)	Community support (percent)
1991	3	60	5
1992	5	19	30
1993	10	25	35
1994	15	30	50
1995	25	52	3
1996	35	36	35
1997	20	17	9
1998	18	58	7
1999	13	28	40
2000	6	45	20

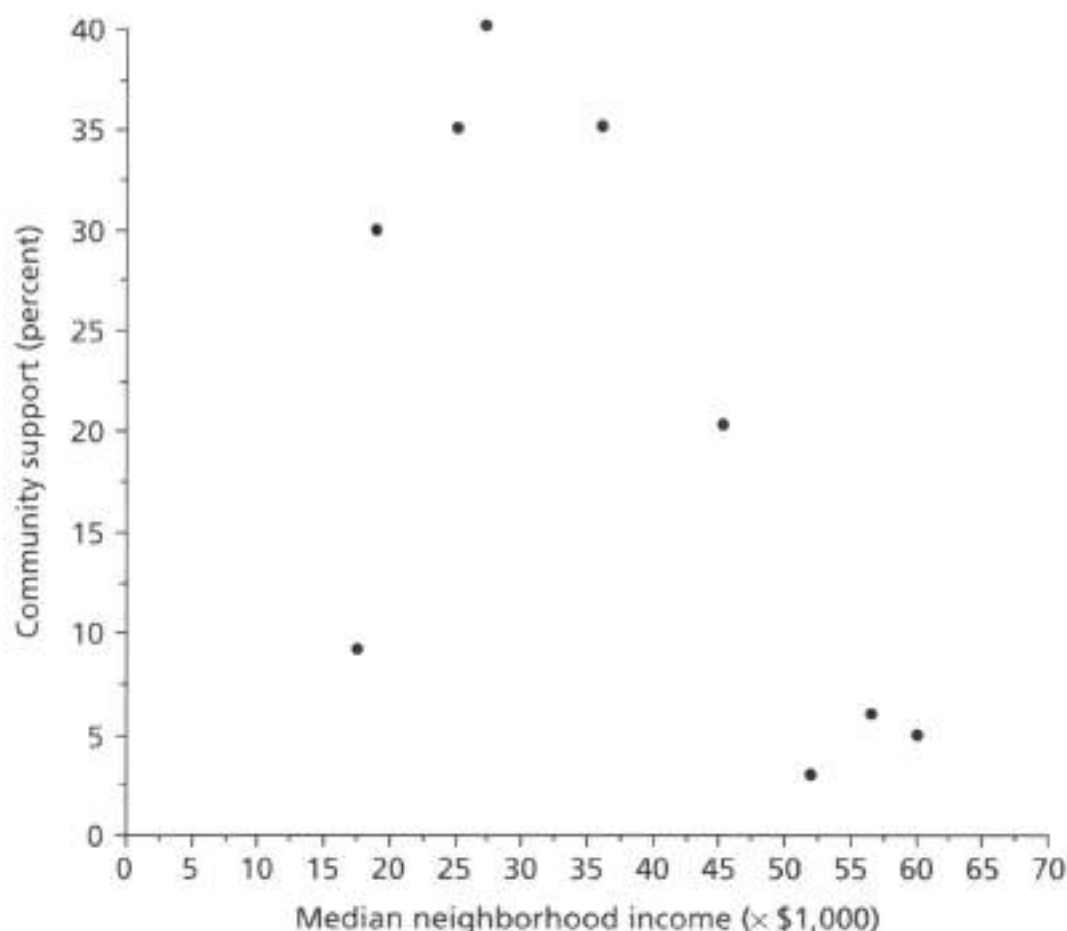
park. The data clearly show that support for the project increased rapidly over the years and then, as the roller-skating fad diminished, support waned. The data are presented in a way (in this case over time) that makes the nature of the relationship between the two variables obvious. However, the results of another series may not be arranged so conveniently for the researcher to discern the shape of the relationship. Take series B, which shows a cross section of support among community residents for the same project in 2001. Since the dependent and independent variables are not arranged in any order, you may not be able to grasp their relationship simply by looking at the data. However, I have plotted the two variables in Figure B.4, which clearly shows a quadratic relationship between them, with support for the park coming solidly from the middle-income neighborhoods. In contrast, poor and wealthy citizens show a distinct lack of interest.

Although Figure B.4 strongly suggests a nonlinear relationship, we have not solved the problem completely. You may still be undecided about the type of polynomial to use. You could try a quadratic or a second-degree polynomial.

The estimation of the second-degree polynomial gives us the following result,

$$\begin{aligned}
 P &= -24.92 + 3.71 Y - 0.056 Y^2 \\
 t &= (-0.92) \quad (2.37)^* \quad (-2.80)^* \\
 R^2 &= 0.69 \\
 \text{Adjusted } R^2 &= 0.61 \\
 F(2, 7) &= 7.91^* \\
 n &= 10
 \end{aligned}
 \tag{B.8}$$

*Significant at the .05 level

FIGURE B.4 Community Support for Roller-skating Park, 2001

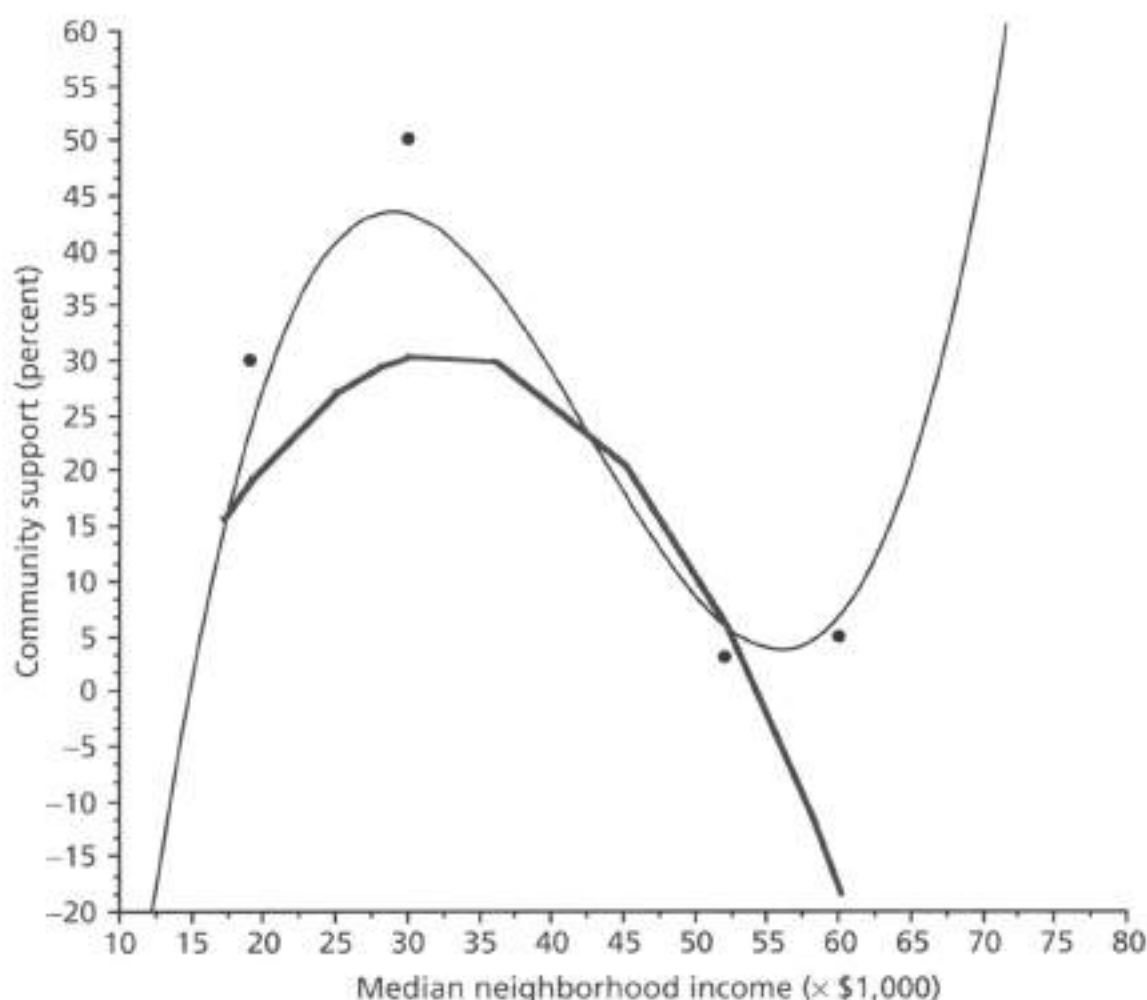
where P is the percentage of support for the project within the community in 2001, and Y is the median income of the community.

The estimated results look fine, with high R^2 and significant t and F values. The estimated curve has been plotted with a heavy line in Figure B.5.

However, we may also be tempted to try a third-degree curve, since the series shows some sign of turning the corner and going up again. In that case, the results are as follows:

$$\begin{aligned}
 P &= -192.78 + 19.69 A - 0.515 Y^2 + 0.004 Y^3 \\
 t &= (4.9)^* \quad (5.45)^* \quad (5.06)^* \quad (4.54)^* \\
 R^2 &= 0.93 \\
 \text{Adjusted } R^2 &= 0.90 \\
 F(3, 6) &= 26.88^* \\
 n &= 10
 \end{aligned}
 \tag{B.9}$$

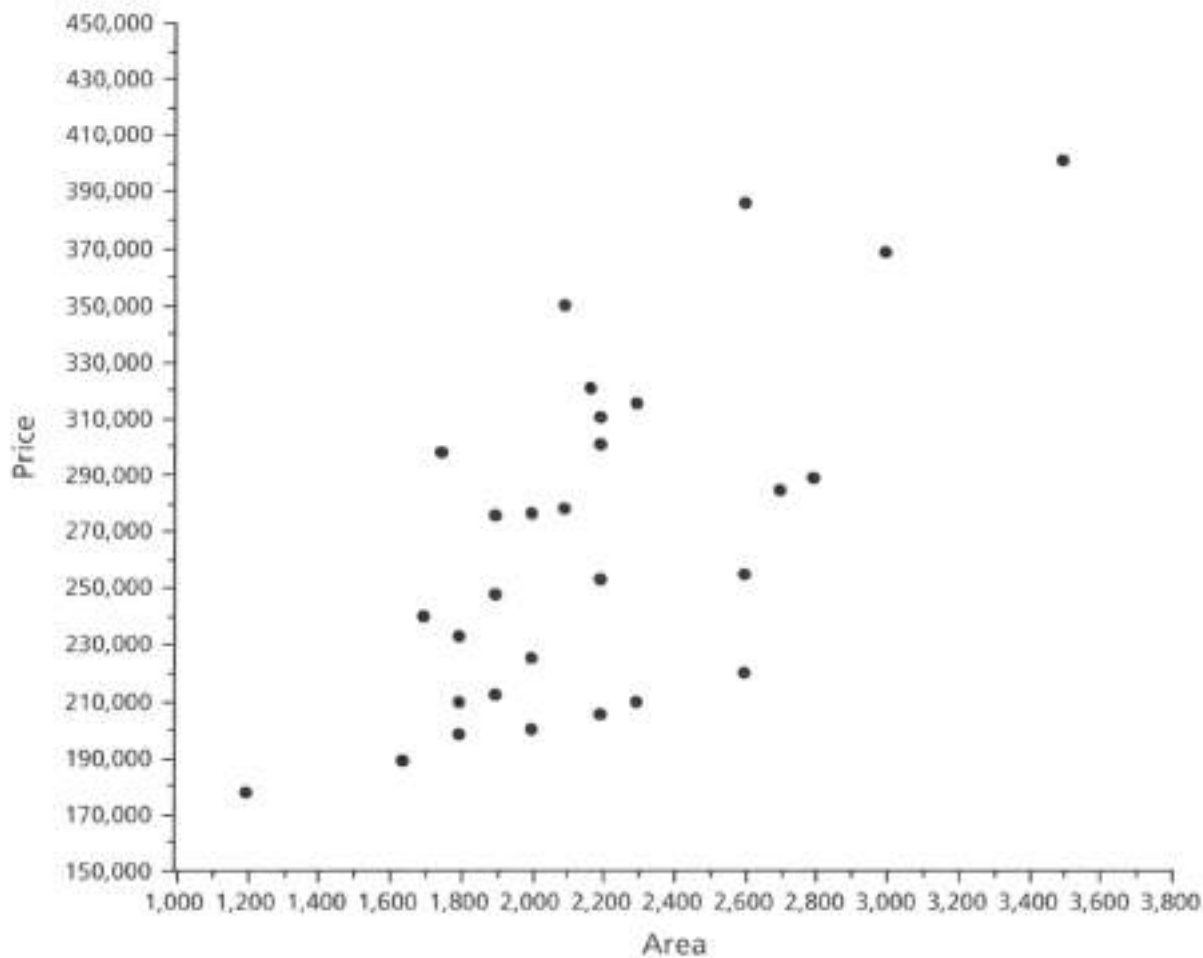
*Significant at the .05 level

FIGURE B.5 Projection and Model Specification

The preceding results indicate that although the second-degree curve gives us a good fit, the third-degree one gives us an even better fit. I have plotted the expected values on the basis of the preceding formulations (equations B.8 and B.9) in Figure B.5. Not only do the two models give us different estimated results for the observations within the sample range, but also the problem of model specification becomes truly critical when we attempt to project outside the sample range. For example, if we want to project for $Y = 75$, the expected value for the quadratic form is -62.42 , while for the third-degree polynomial, it is 74.6 .⁴

It is important to remember that the functional form you choose may give you forecasts that are significantly different from one another, especially for the independent variables *outside the sample range considered in the model* (the range between the highest and the lowest values of the independent variable). In that case, you may not have any objective way of choosing among the different functional forms.

However, even if a higher order functional form gives you an apparently good fit, there are few, if any, instances in the social sciences that conform to the explosive

FIGURE B.6**Relationship between Housing Prices and Covered Area**

growth of a higher order polynomial. Therefore, to repeat my previous suggestion, always use a linear form when in doubt, and never use a functional form higher than the second order.

Returning to our example of predicting housing prices, to specify the functional form, price is plotted against area in Figure B.6. From the plotted data, we decided that the linear form was the best functional form.

WHEN REGRESSION RESULTS ARE SUSPECT: THE ERRORS OF ESTIMATION

When using regression analyses, a researcher, like the heroes of the Greek epics, can be misled by a number of different errors or may forfeit sound reasoning and be seduced by empirical results that appear to be gorgeous at first glance. However, more careful scrutiny reveals their deceptive nature. Let us discuss the major errors of regression analysis.

Identifying Multicollinearity

One of the premises of the classical least squares method is that the independent variables are not *collinear* (not correlated). In statistical terms this lack of correlation is called *orthogonal*. Take any series of numbers (for example, the square footage measurements of the various houses in our example). If another variable is created by adding, multiplying, dividing, or subtracting a constant, these two variables are perfectly correlated and are called perfectly collinear with each other. Hence, in our housing price equation, if we add another variable that measures the covered area in square meters, we have created a situation of perfect multicollinearity, since this new variable is a multiple of the old variable. The presence of perfect multicollinearity prevents us from deriving any estimate of the slope coefficients.⁵ Let me explain why. Suppose we are trying to estimate the following equation

$$Y = a + bX + cZ + \varepsilon \quad (\text{B.10})$$

where X and Z are two highly correlated (collinear) independent variables.

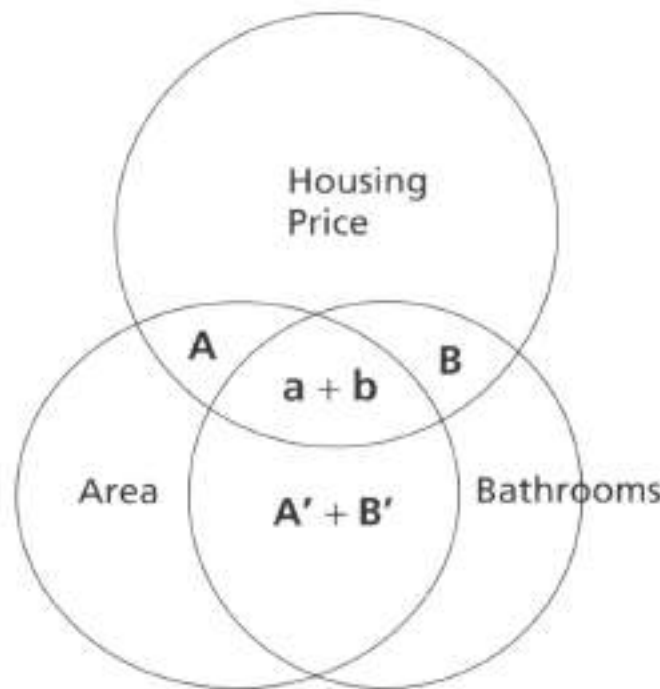
You may visualize the effects of multicollinearity with the help of the Venn diagram in Figure B.7.

Let us go back to our example of estimating housing price. Suppose that we have included two independent variables: the square footage of the house (area) and the number of bathrooms (bathrooms). As you can see from this diagram, the two independent variables explain the dependent variable. The parts of housing price that are explained by the two independent variables are the overlapping areas. However, if you compare this figure with Figure B.2, you will see that in this case, the problem is that the two independent variables overlap to a large extent. This result is caused by the simple fact that the larger a house is, the more bathrooms it is likely to have. This is the problem of multicollinearity.

To explain the problem, I have marked the areas of intersection separately. The parts of housing price that are explained exclusively by each of the two independent variables, area and bathrooms, are labeled A and B , respectively. The segment that is jointly explained is marked $a + b$. Finally, the portion of the figure that is a simple overlap between the two independent variables, without any implication for the explanation of housing price, is labeled $A' + B'$. As we have seen before, R^2 is measured by the ratio $[A + (a + b) + B]/(\text{total area of housing price})$.

The problem of multicollinearity shows up in the estimation of the coefficients. If there is a high degree of multicollinearity among the independent variables, we encounter two significant problems. First, we face the problem of estimating the "true" coefficients of the individual independent variables. Thus, in Figure B.7, while calculating the relation between housing price and area, regression analysis concentrates only on segment A . Because it cannot disentangle the area $(a + b)$ jointly explained by the two independent variables, the regression model ignores area $(a + b)$. As a result, *the estimated coefficients turn out to be smaller and more statistically insignificant than they would be if only one of these variables were introduced without the other.*

FIGURE B.7

Multicollinearity and Its Effects on R^2 

Second, multicollinearity enlarges the standard error of the estimated coefficients. As a result, the calculated t values can become highly questionable (see Figure B.8).

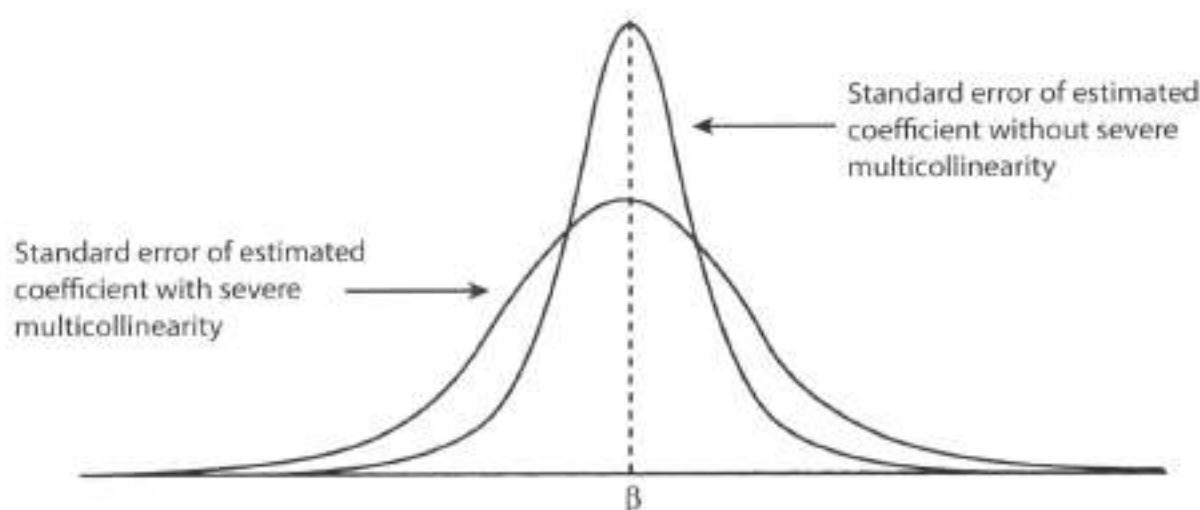
Computing a multiple regression is extremely time-consuming, so the formulas for calculating regression coefficients for multiple regression models are not included in this book. However, for the purposes of exposition, I would like to point out that when there is more than one independent variable, the estimated standard error of slope coefficients is obtained with the following formula:

$$SE(b) = \sqrt{\frac{\sum_{i=1}^n e_i^2 / (n - k)}{\sum_{i=1}^n (X_i - \bar{X})(1 - r_{XZ}^2)}} \quad (\text{B.11})$$

where e_i^2 is the error sum of squares, n is the number of observations, k is the number of independent variables, and r_{XZ}^2 is the square of the correlation coefficient between the two independent variables, X and Z .

From the preceding formula, you can see that when there is no correlation between X and Z (that is, when $r_{XZ}^2 = 0$), the term $(1 - r_{XZ}^2)$ in the denominator is equal to 1. Therefore, the estimated standard error of the estimated regression coefficient b is at its minimum.

However, as the correlation between X and Z becomes stronger, the correlation coefficient, r_{XZ}^2 , approaches 1. If there is a perfect correlation, this term equals 1. In that case, the denominator of the estimated standard error equals 0, which makes

FIGURE B.8**The Effects of Multicollinearity on the Standard Errors of Estimation**

the standard error infinite. This result makes the estimation of the standard error of the regression coefficients impossible.

In real life we seldom encounter perfect collinearity between two independent variables unless they are the same variable, simply expressed in two different measurement units. If we accidentally use two perfectly collinear independent variables, our computer software will warn us by posting an error message. What the software cannot tell us, though, is where there is less than perfect correlation. This situation, as we can deduce from the preceding equation, will make our estimated standard error more widely dispersed, as shown in Figure B.8.

What is the immediate effect of this increasing dispersion of the error term of the estimated coefficients? As you may recall, the t values for the estimated coefficients are calculated by dividing the coefficient by its standard error (see equation B.11). Therefore, this increase in the error term will reduce the t value, causing the researcher to reject an otherwise significant independent variable as insignificant.

You may note that despite multicollinearity, the estimated coefficients will remain unbiased, as Figure B.8 indicates. In other words, in our example of housing prices, we estimated the regression coefficients with the help of thirty observations. This calculation gave us an estimate of the true value of b . If we considered another sample group of houses and reestimated the same equation, we would get another estimate of this elusive b . If we continued to repeat this experiment, the distribution of the estimated values of b would have a mean, which would get closer and closer to the true value of b . You may recall that this is the property of unbiasedness in the classical least squares method. The presence of multicollinearity does not make our estimated coefficients biased. Therefore, this property of least squares is not violated.

Since the variance of the distribution of the estimated coefficients goes up with increasing levels of multicollinearity, including a collinear variable will cause the estimates to fluctuate widely. If we include an independent variable that is highly correlated with another independent variable in the model, the variables' estimated coefficients

will become extremely sensitive and will change dramatically each time the equation is run with a slightly different set of independent variables. *However, this extreme sensitivity will not affect those independent variables that are not highly correlated with this new variable.*

The presence of multicollinearity does not affect the measures of overall goodness of fit. That is, R^2 and F values in an equation with a high degree of multicollinearity will be unaffected. Consider our example of estimating home prices. Suppose we include two other variables in the model, the number of bathrooms (bathrooms) and the total, as opposed to covered, land area (land). Suppose that these homes are all located in the same tract-home development, and therefore the total land area is highly correlated with the covered area. So is the number of bathrooms. These new variables are presented in Table B.4.

TABLE B.4 Housing Prices and Multicollinearity

Price (\$)	Area (square feet)	Bathrooms	View	Cul-de-sac	Land (square feet)
310,000	2,200	3.5	1	1	4,170
233,000	1,800	2.5	0	0	3,330
400,000	3,500	5.0	1	0	6,175
430,000	3,200	4.0	1	1	6,000
210,000	1,800	2.0	0	0	3,330
240,000	1,700	2.0	1	1	3,045
300,000	2,200	3.0	0	1	4,090
350,000	2,100	3.0	1	1	3,885
385,000	2,600	4.5	1	1	4,810
368,000	3,000	4.5	0	0	5,650
200,000	2,000	2.5	0	0	3,777
298,000	1,750	2.0	1	1	3,237
275,000	1,900	2.5	0	1	3,500
198,000	1,800	1.5	0	0	3,330
253,000	2,200	3.0	1	0	4,100
278,000	2,100	3.0	1	1	3,885
320,000	2,170	3.5	1	1	3,999
178,000	1,200	1.0	1	1	2,220
225,000	2,000	3.0	0	0	3,700
212,000	1,900	2.0	0	1	3,515
288,000	2,800	3.5	0	1	5,280
315,000	2,300	4.0	1	1	4,255
255,000	2,600	3.0	0	1	4,810
284,000	2,700	3.5	1	0	4,995
189,000	1,640	2.0	1	0	3,034
220,000	2,600	3.0	0	0	4,810
248,000	1,900	2.5	1	0	3,515
276,000	2,000	3.5	1	1	3,700
210,000	2,300	3.0	0	0	4,255
205,000	2,200	2.5	0	0	4,070

Let us explore the effects of including these highly correlated independent variables in the model. The results of the newly estimated equation are as follows:

$$\begin{aligned}
 \text{Price} = & 33,268.056 + 29.92 \text{ Area} + 30,137 \text{ View} + 38,598.56 \text{ Cul-de-sac} \\
 & + 10.24 \text{ Land} + 31,879.62 \text{ Bathrooms} \\
 t = & (1.10) \quad (0.19) \quad (2.24)^* \quad (3.15)^* \quad (0.12)^* \quad (2.29) \\
 R^2 = & 0.825 \\
 \text{Adjusted } R^2 = & 0.788 \\
 F(5, 24) = & 22.58^* \\
 n = & 30
 \end{aligned}
 \tag{B.12}$$

*Significant at the .05 level

By comparing the results of equation B.12 with those of equation B.3, we can see the impact of severe multicollinearity. The most affected variable is area, which is highly correlated with the newly introduced variables, land and bathrooms. Although the estimated coefficients and their respective t values for the other two independent variables (view and cul-de-sac) remained relatively unchanged, the estimated coefficient for area was significantly reduced, and its error term greatly increased, causing a precipitous drop in the t value. You also may notice that the R^2 value was not affected by including the highly correlated independent variables.

Once you understand the problems caused by multicollinearity, the question becomes how to detect it and what to do about it. You should note that multicollinearity should be suspected if you have high R^2 values but low t values.

To detect which independent variables are collinear, you should calculate the correlation matrix among the independent variables. Most software packages can provide you with the matrix of correlation coefficients shown in Table B.5.

Each entry in the table shows the correlation coefficient between the row and column variables. For example, the correlation between bathrooms and area is .867. The diagonal numbers are all 1.000 because, by definition, each variable has a perfect correlation with itself. You can see that three of the independent variables—area, bathrooms, and land—are highly correlated with each other. This would account for the multicollinearity in the estimated equation.⁶

TABLE B.5 Correlation Matrix

	Area	Bathrooms	View	Cul-de-sac	Time	Land
Area	1.000					
Bathrooms	.867	1.000				
View	-.003	.208	1.000			
Cul-de-sac	-.078	.060	.330	1.000		
Time	-.079	.035	.289	.237	1.000	
Land	.997	.864	-.022	-.069	-.092	1.000

Resolving Multicollinearity

Multicollinearity poses a dilemma for the analyst. If there is severe multicollinearity between two independent variables, one of the best ways to resolve the problem is to eliminate one of the variables. However, if the eliminated variable happens to be an important variable in explaining the dependent variable, then by eliminating it we will cause specification error with its accompanying problems. Furthermore, as we have noted, multicollinearity among a partial list of independent variables will not affect the other variables, nor will the error term be biased.

Researchers often choose not to do anything about multicollinearity unless the problem is acute. Looking at the correlation coefficients of our previous example, I would be inclined to include the number of bathrooms in the model because it happens to be one of the vital considerations in determining the price of a home. On the other hand, I would reason that with a correlation coefficient of .997, land is not adding much to the explanation of market price beyond that provided by area. I would specify the model by including bathrooms but eliminating land. As you can tell, this decision is purely a line call on the basis of subjective judgment.

Econometricians have tried to walk the tightrope between being comprehensive—including all the important information—and avoiding severe multicollinearity. Although balancing these competing objectives is a matter of subjective judgment and comes through years of practice, you can improve your chances of success immediately by following these practical steps for eliminating multicollinearity among the independent variables:

1. *In certain circumstances the two independent variables can be added together.* This newly created variable will contain information from both variables and may add to the explanation without contributing the problems of multicollinearity. This trick may work provided the two variables do not have opposite expected signs or are not significantly different in magnitude. Suppose we are adding two highly correlated independent variables, X and Z , to form a new variable. If X is positively related to the dependent variable, Y , while Z is negatively related (X and Z are negatively correlated with each other), the newly created variable will have little explanatory capability, because the positive relation will be offset by the negative one. Adding independent variables also will not work if one of the variables has a substantially higher mean than do the others. In this case the smaller variable will be lost in the larger variable, and the linear combination of the two will not provide any more insight into the variation of the dependent variable. Furthermore, the newly created variable may not have any intuitive meaning. For example, for a completely different purpose, economist Arthur Okun created an index by adding the unemployment rate and the rate of inflation, calling the sum the “misery index.” This composite measurement has a readily understandable meaning. In contrast, if we add area and land, this composite variable may not convey any definite meaning to policymakers (in this case, the property owners).

2. *For time-series data, the problem of multicollinearity can also be solved if you take the first difference.* That is, you create a new variable by subtracting the preceding time period’s data from those of the current period. Suppose we are trying to forecast the sales tax revenue for a state government. We have chosen as independent variables the growth rate of per capita state income and the rate of unemployment. But in formulating our model, we

find that there is a strong negative correlation between these variables (that is, during prosperous times, when the income growth rate is high, unemployment is low, and vice versa). Therefore, we cannot add these two variables together. A way out may be to use the yearly difference in the unemployment rate, which may not have as strong a correlation with the growth rate of per capita income as does the absolute level of unemployment. However, even this method is not a panacea for correcting multicollinearity. Similar to the problem of creating a composite variable, using the yearly difference of a variable may not have the same meaning (or even any meaning) as an independent variable. Even after getting high t and R^2 values, we may be at a loss to explain our results in a meaningful way.

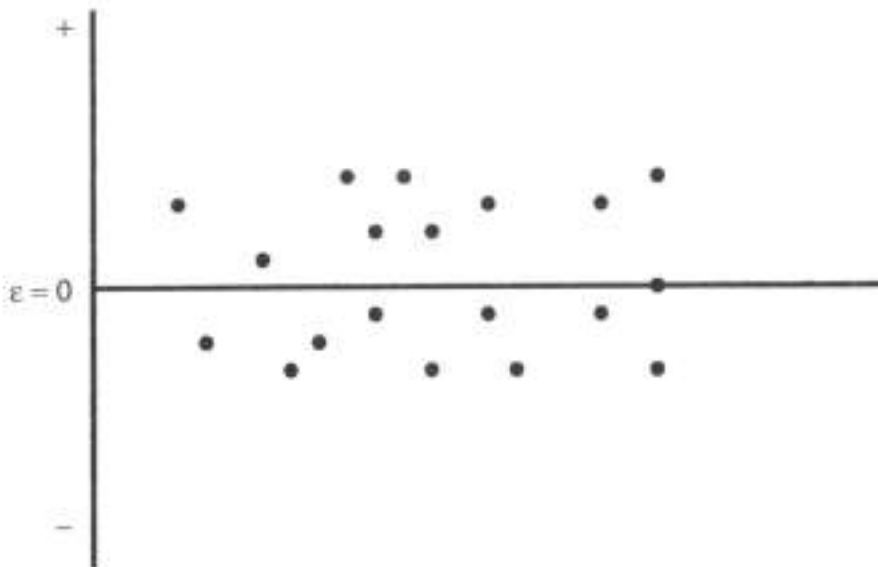
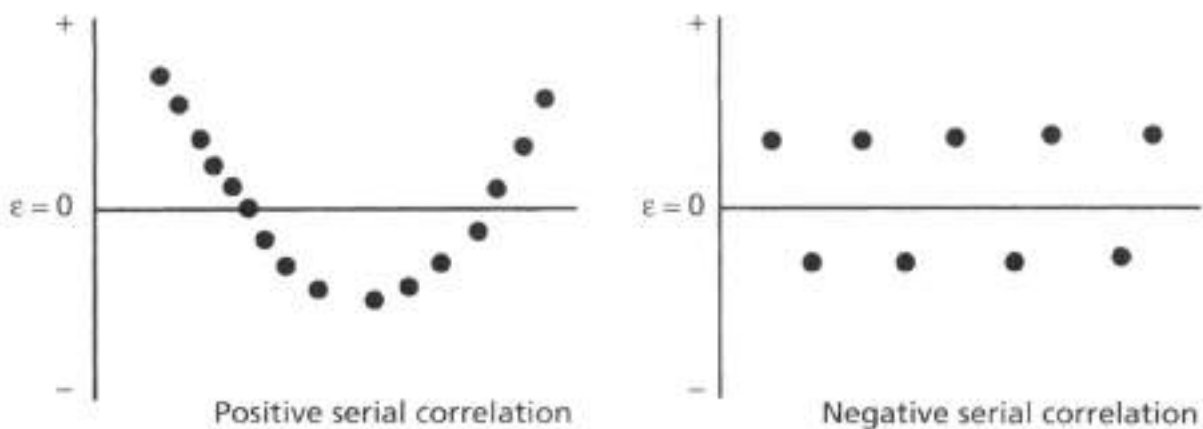
3. *Finally, another (and probably the least controversial) way of dealing with multicollinearity is to increase the sample size.* If doing so is a viable alternative, it is certainly worth pursuing. If the two variables are not simply multiples of each other (perfectly correlated), then as the number of observations increases, the natural variations within the series will sufficiently distinguish themselves and allow the proper estimation of the model. In our example of housing prices, if we can increase the sample size from thirty to, say, 300, the variations within the two series may distinguish themselves enough to allow us to estimate the model properly.

Identifying Serial Correlation

Serial correlation (also known as autocorrelation) means that the order in which the observations are arranged has some effect on the estimation of the regression coefficients. In other words, serial correlation exists if the error of one observation depends on that of the previous one. An example may clarify the point. Many cities across the nation want to reduce lawsuits stemming from work-related injuries. The job of a risk manager is to curtail such lawsuits by offering adequate training to his or her workers. However, the training programs are expensive, so some cities establish benchmarks for the amount of workers' compensation paid out each year. If a city pays more than the specified amount this year, it will invest more money in the training program next year. On the other hand, if the city pays less than this benchmark amount, less money will be allocated for training the following year. Before initiating this benchmark program, the allocation of money followed other guidelines, and the error term from the equation explaining expenditure on training showed a random pattern. I have shown the hypothetical distribution of the error term in Figure B.9, in which the trend line over time is a flat horizontal line.

However, after introducing the benchmark program, the error terms were correlated with the past year's term. That is, if the yearly compensation overshoot the target, more money flowed into the training program; if it undershot, money was taken away from the program. Situations such as these, in which the current year's performance is predicated upon that of the previous year, show serial correlation. In contrast to Figure B.9, in which the error terms are randomly distributed, Figure B.10 shows two kinds of serial correlation: positive and negative. If the current period's error term generally shows the same sign as error terms of previous periods, the series is said to have a positive serial correlation. Thus, if a public policy affects the course of the economy for a number of years (as did Proposition 13, the California initiative that cut property taxes and restricted the ability of local governments to raise and spend money), the series shows positive serial correlation.

In contrast, if a series shows alternately positive and negative error terms, it demonstrates negative serial correlation. In this case the error terms oscillate back and forth

FIGURE B.9 Distribution of Errors Showing No Serial Correlation**FIGURE B.10** Serial Correlation

like a pendulum. The equation explaining spending on the benchmark program may exhibit a negative serial correlation.

A correlation between successive error terms violates one of the fundamental principles of the least squares method. The most widely used method for detecting serial correlation is the Durbin-Watson d test.⁷ This test applies only in the following cases:

1. The regression model includes an intercept term.
2. The error terms have a first-order correlation, meaning that each year's error is correlated only with that of the preceding year and not with the errors of the previous two or three years. In symbolic terms this can be written as

$$e_t = \rho e_{t-1} + \varepsilon \quad (\text{B.13})$$

where e_t is the residual (estimated error term) in year t , e_{t-1} is the residual (estimated error term) in the previous year ($t-1$), ρ is the correlation coefficient, and ε is the true random error term.

3. The regression model does not include a lagged dependent term.

The Durbin-Watson d statistic is defined as

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad (\text{B.14})$$

The d statistic varies between 4 and 0. If there is perfect positive serial correlation ($r = 1$), e_t is the same as e_{t-1} , and the numerator becomes 0. In other words, with a perfect positive correlation, the d statistic equals 0. If, on the other hand, there is perfect negative correlation ($r = -1$), the numerator becomes $4\sum e_t^2$. In that case equation B.14 becomes

$$\frac{4\sum e_t^2}{\sum e_t^2} = 4$$

If there is no serial correlation at all, the d statistic is equal to 2. Hence, if the d statistic is close to either 0 or 4, you should suspect serial correlation.

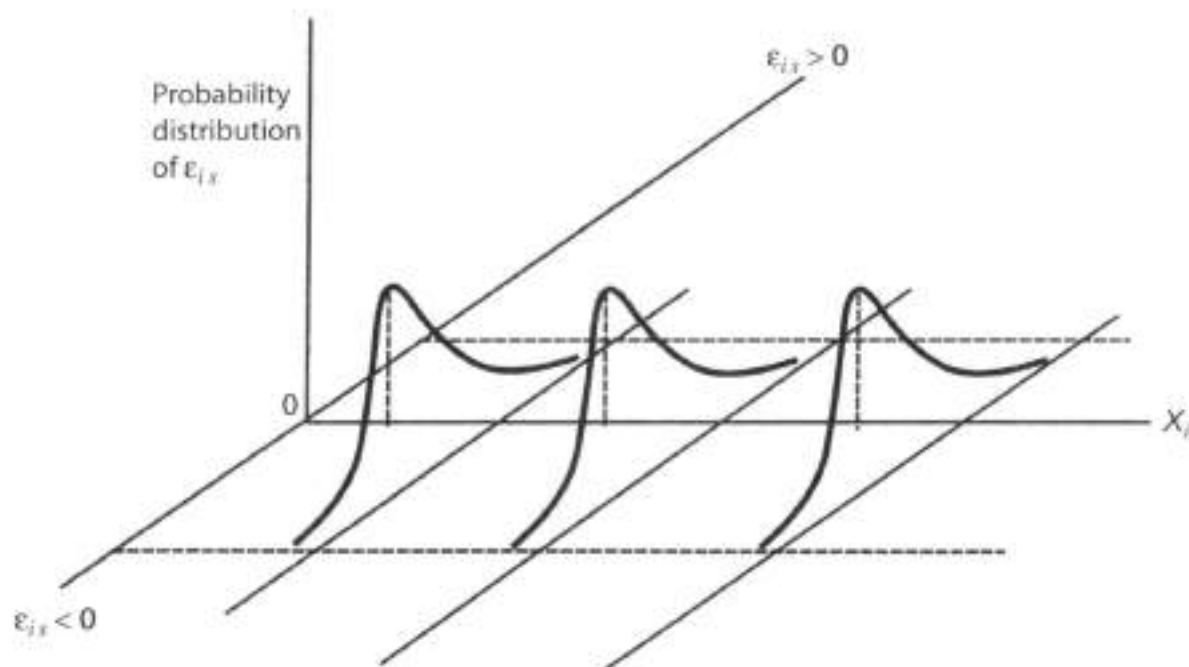
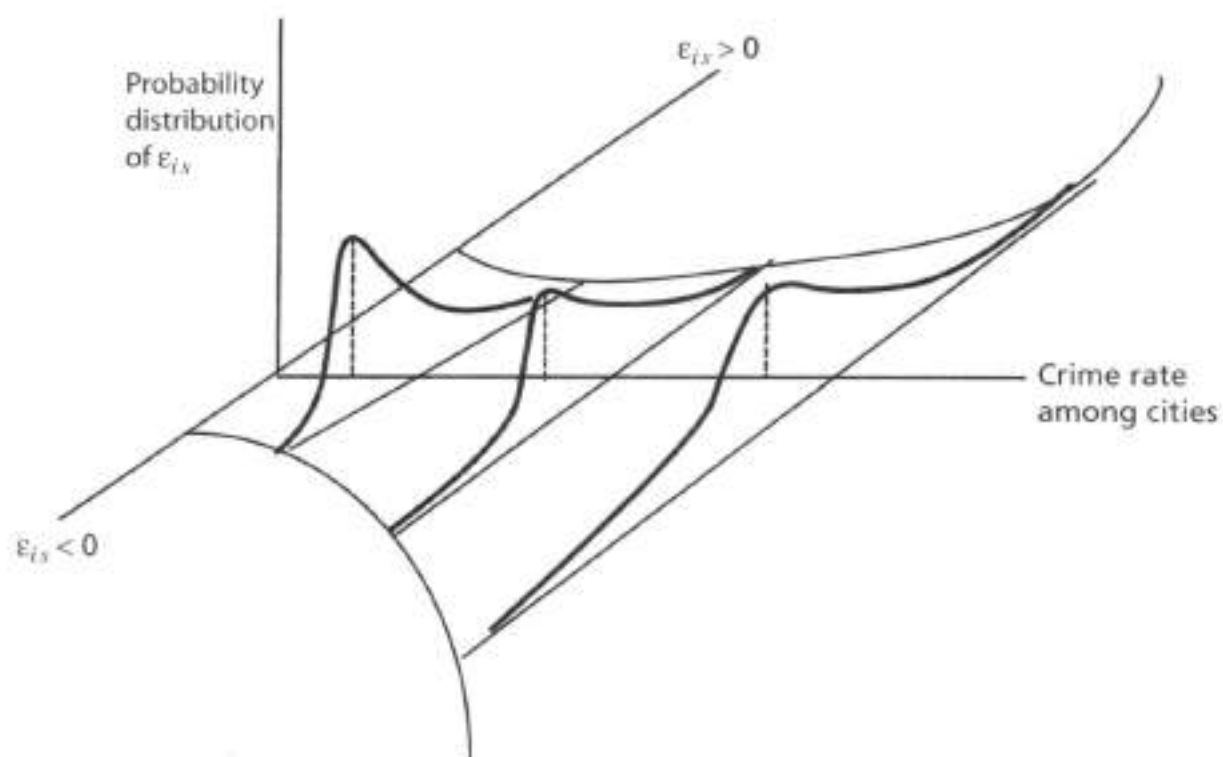
Unfortunately, the correction for serial correlation is beyond the scope of this introductory textbook. If you are interested in knowing more about these problems, you may consult one of many excellent texts on econometrics.⁸

Heteroskedasticity: The Problem of Scaling Variables

One of the important conditions for the least squares method is that the variability (standard deviation) of the error term of the observations does not vary with the size of the dependent variable. This is the condition of homoskedasticity. Figure B.11 demonstrates the implication of this assumption.

As you can see, the variance of the error term corresponding to the three observations of the dependent variable X falls within the same band. However, this happy situation may not hold true if the independent variables vary a great deal with the size of the dependent variable. In that case, as the scale of the dependent variable increases, so does the variability of the error term. This situation of increasing variability of the error term is called heteroskedasticity and is shown in Figure B.12.

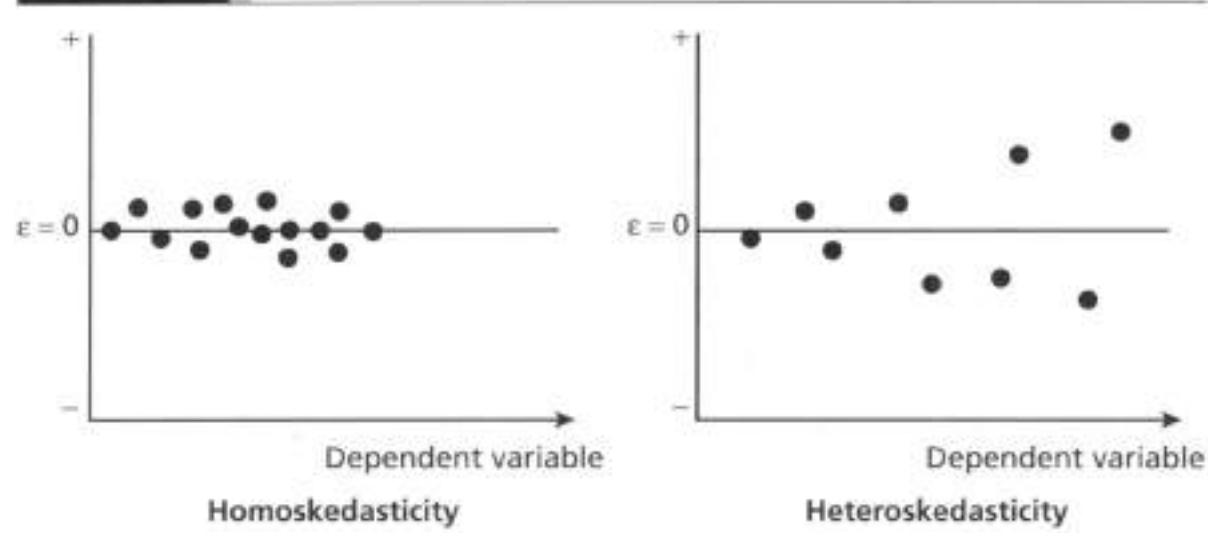
Heteroskedasticity, therefore, is caused essentially by mixing apples with oranges in the data set and is typically encountered in cross-section models. Suppose we are trying to account for urban crime in America. If our data set contains the number of high crimes committed in large cities, along with those perpetrated in small towns in primarily agricultural states, we are likely to encounter heteroskedasticity. If the model is not properly specified, differences in the variance of the error terms may result. This is called impure heteroskedasticity. If, on the other hand, despite the best specification of the model, the error terms show signs of heteroskedasticity, this is known as pure heteroskedasticity.

FIGURE B.11 Homoskedastic Distribution of Errors**FIGURE B.12** Heteroskedastic Distribution of Errors

Detecting Heteroskedasticity

Heteroskedasticity can be visually inspected by plotting the error terms. If we do not have heteroskedasticity, the errors will be distributed randomly around zero. However, if it is present, the errors will show a flaring-out pattern, as in Figure B.13.

FIGURE B.13 Hypothetical Errors Showing Presence and Absence of Heteroskedasticity



To explain the point, let us return to our example of estimating housing prices. Using the results of equation B.3, we have obtained the predicted value for each of the homes in the sample. The difference between the actual and the predicted value is the error term. The error terms of the housing price data have been plotted in Figure B.14.

FIGURE B.14 Error Terms of Housing Price Data

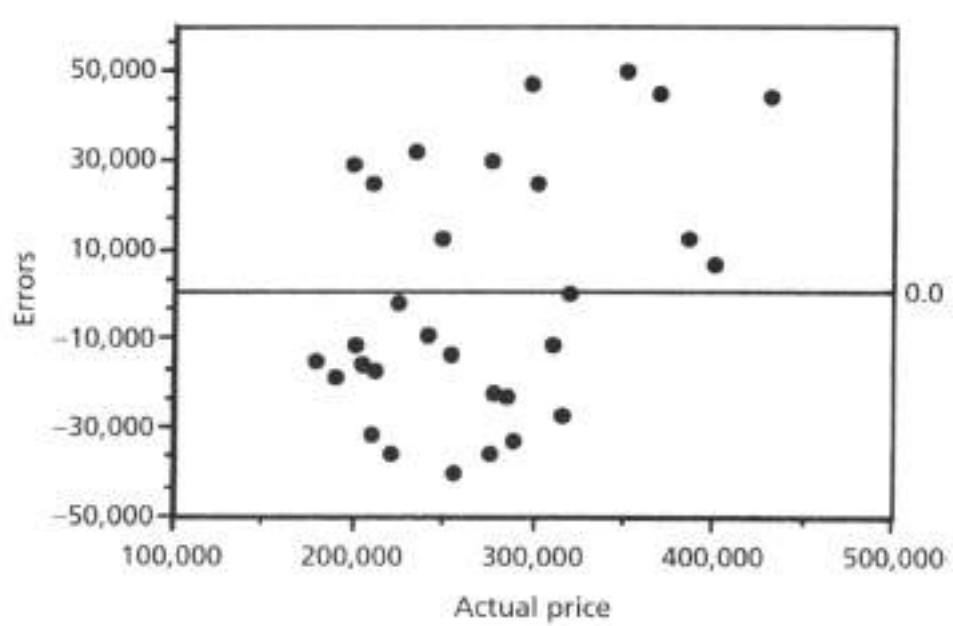


TABLE B.6

Housing Prices with Data for a Dissimilar Neighborhood

Observation	Price (\$)	Predicted price (\$)	Error (residuals)
1	310,000	320,733	-10,733
2	233,000	168,067	64,933
3	400,000	422,349	-22,349
4	430,000	436,803	-6,803
5	210,000	168,067	41,933
6	240,000	262,698	-22,698
7	300,000	263,770	36,230
8	350,000	309,126	40,874
9	385,000	367,161	17,839
10	368,000	307,351	60,649
11	200,000	191,281	8,719
12	298,000	268,501	29,498
13	275,000	228,949	46,051
14	198,000	168,067	29,933
15	253,000	271,458	-18,458
16	278,000	309,126	-31,126
17	320,000	317,251	2,749
18	178,000	204,663	-26,663
19	225,000	191,281	33,719
20	212,000	228,949	-16,949
21	288,000	333,412	-45,412
22	315,000	332,340	-17,340
23	255,000	310,198	-55,198
24	284,000	329,493	45,493
25	189,000	206,458	-17,459
26	220,000	260,923	-40,923
27	248,000	236,637	11,363
28	276,000	297,519	-21,519
29	210,000	226,102	-16,102
30	205,000	214,495	-9,495
31	535,000	548,433	-13,433
32	648,000	640,217	7,783
33	656,000	606,468	49,532
34	802,000	814,322	-12,322
35	735,000	724,313	10,687
36	546,000	658,699	-112,699
37	762,900	944,846	-181,946
38	485,000	546,637	-61,637
39	942,000	979,667	-37,667
40	1,250,000	898,418	351,582

The plot in Figure B.14 does not indicate overwhelming heteroskedasticity, since the housing prices are fairly close to one another. Now suppose we included in the sample data an adjoining but dissimilar neighborhood. Whereas the original sample contains data from an upper-middle-class area, the next sample is from a decidedly more affluent one.

The expanded data set, with ten new entries of extremely high-priced homes, is presented in Table B.6. In this table, we have included a column of predicted values of house

prices by using the specification of equation B.15 and adding a dummy variable for the two communities. The value of this newly created dummy variable is 0 for the moderately priced community (Comm) is 0 and 1 for the wealthy community. The estimated equation is as follows:

$$\begin{aligned} \text{Price} = & -40,859.71 + 116.07 \text{ Area} + 56,963.41 \text{ View} + 49,275 \text{ Cul-de-sac} \\ & + 275,903.98 \text{ Comm} \\ t = & (0.94) \quad (6.20)^* \quad (2.11)^* \quad (1.88)^* \quad (6.73)^* \end{aligned} \quad (\text{B.15})$$

$$R^2 = 0.905$$

$$\text{Adjusted } R^2 = 0.894$$

$$F(4, 35) = 83.57^*$$

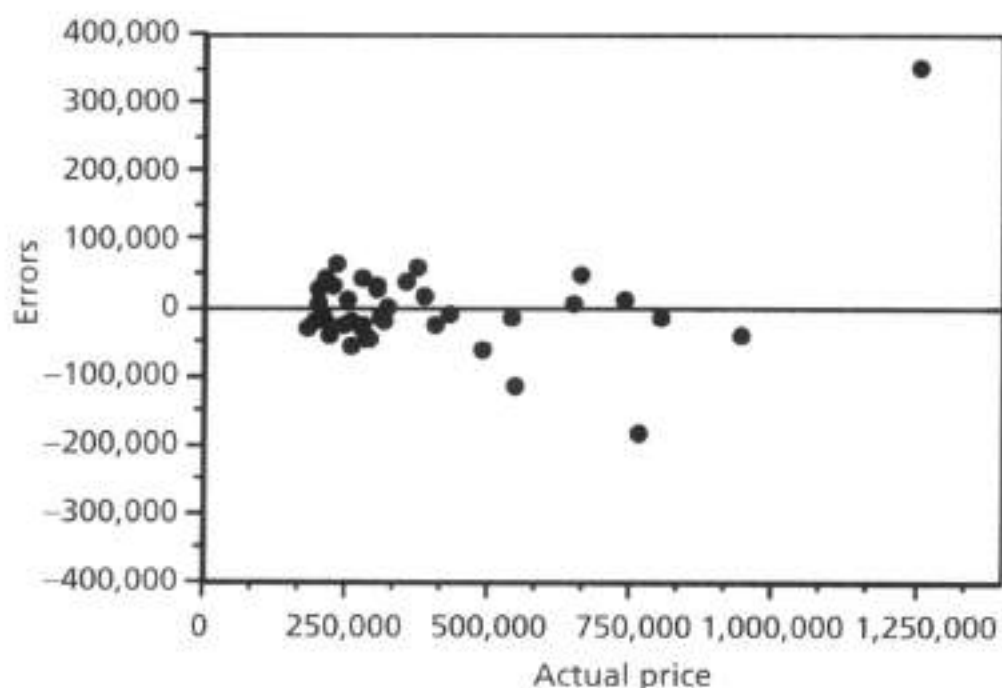
$$n = 40$$

*Significant at the .05 level

The column "Predicted price" was created using this formula. The error terms were calculated by subtracting the predicted price from the actual price (the price variable).

The terms shown in the "Error" column of Table B.6 are plotted against price in Figure B.15. By comparing this figure with the previous one (Figure B.14), you can clearly see that the introduction to the sample of high-priced homes has created heteroskedasticity.

FIGURE B.15 Error Terms Showing Heteroskedasticity



Effects of Heteroskedasticity

Pure heteroskedasticity can increase the variance of the estimated coefficients in a way that will cause the t and F values to come out stronger than what is warranted by the data. Therefore, though the estimates will remain unbiased, all the measures of hypothesis testing will be suspect. On the other hand, if there is impure heteroskedasticity resulting from a poor specification of the equation, its impact on the estimators will be similar to the effect of serial correlation discussed earlier.

Correcting Heteroskedasticity

The error of heteroskedasticity is primarily caused by a variable that is too diverse. For that reason, heteroskedasticity should be suspected in cross-section data with samples varying greatly in size. Time-series data stretching over a long period of time are also suspect, as are data series in which the magnitude of the dependent variable increases with time at a very high rate. If we are trying to predict the number of homicides in our city on the basis of the number of homicides committed this year in a large number of diverse cities and towns, we are likely to encounter heteroskedasticity. Also, it is likely to exist in time-series data showing the number of individuals with AIDS, since the actual number of individuals with the disease keeps increasing every year.

Heteroskedasticity can be detected with the help of some of the more sophisticated tests, which are beyond the scope of this introductory book. However, we may say a few words about correcting heteroskedasticity at this point. First, if heteroskedasticity is impure, it can be corrected simply by better specifying the model. To repeat our discussion of proper specification, developing a better model requires a great deal of hard work in thinking through the proper theoretical structure.

If, on the other hand, heteroskedasticity is pure, the very nature of the data causes the variance of the error terms to change, and we need to think about the dependent variable. For example, if we encounter strong heteroskedasticity in trying to work with data on the number of homicides in cities and towns across the nation, we may separate the sample into two groups: large cities and small towns. Or we may express the absolute numbers in ratios, such as homicide rate per 1,000 inhabitants. Sometimes this kind of data transformation can reduce the variance to such an extent that heteroskedasticity can be significantly reduced.

When the Data Are Imperfect

Finally, in forecasting with regression models, serious errors occur when the data are spurious. In statistics this situation is known as the error in variable, or the measurement problem. The problem can creep in because of various factors. Many sociological data, when available, often contain severe measurement errors. Thus, information on child and spousal abuse, rape, and incest are significantly underrepresented. Even economic data, such as unemployment or inflation figures, are often criticized for containing biases of all sorts. Unemployment data are typically collected by adding the numbers of all those who registered at unemployment offices. This list, then, does not include those whose unemployment benefits have run out, students or homemakers seeking employment for the first time, or the underemployed (those who are working at levels

far below their capabilities, such as an engineer working as a waiter in a restaurant). The data on inflation are computed by comparing the prices of items in a basket of goods a “typical” consumer would buy. Because of wide differences in consumption habits, the inflation figures published by the Bureau of Labor Statistics are likely to affect individuals differently.

Therefore, when data containing biases are used in a regression model, the accuracy of statistical explanation and prediction is affected. Errors can occur if the dependent variable is biased or if there is measurement error in the independent variable as well. If the dependent variable contains measurement error but the independent variable is generally free of it, and the bias is random, then the measurement error will increase the errors of prediction, while the estimated coefficients remain unbiased. Thus, the predicted results will have more variation, although the estimated coefficients will be unbiased. In other words, if we were to repeat the experiment a sufficient number of times with different samples, the average of all the estimated coefficients would approach the actual coefficient. Theoretically at least, this is a lesser problem than the one caused by an error in the independent variable.

If the independent variables contain measurement biases, bias spreads through the error term (ϵ) in the estimated equation, and this problem cannot be corrected easily. In effect, this bias in the independent variable will cause the independent variable to be correlated with the error term (ϵ).⁹

Beta Coefficient: Measuring the Relative Strength of the Independent Variables

The regression coefficients measure the slope of the independent variable in explaining the dependent variable. However, the size of the coefficient depends on the relative size of the independent variable. If you use actual years as independent variables (for example, 2001, 2002, 2003, and so on), the estimated coefficient will be much smaller than if you use time (1, 2, 3, and so forth). This change does not affect the R^2 value or the t and F values. It simply adjusts the absolute size of the coefficient to fit the relationship. Thus, when we estimate a quadratic relationship ($Y = a + bX + cX^2$), the estimated coefficient c turns out to be a lot smaller than b .

This dependence of the coefficients on the size of the independent variables may pose a problem for someone interested in finding out the relative strength of the independent variables in explaining the dependent variable. One way to determine this relative strength is to standardize the variables. You can standardize a variable by calculating

$$SD(X_i) = \frac{X_i - \bar{X}}{s} \quad (\text{B.16})$$

where $SD(X_i)$ is the standardized value of the variable X , $\bar{X}$ is the mean of X_i , and s is the standard deviation of X_i .

Most statistical software packages provide you with the β coefficient along with the estimated coefficients of the nonstandardized variables.

SUMMARY: STEP-BY-STEP SUGGESTIONS FOR BUILDING A MODEL OF CAUSAL PREDICTION

1. **Develop a theory.** The first step in building a causal model is to have an excellent understanding of causality. Do a thorough job of reading the existing literature on the issues. Then, on the basis of your theoretical understanding of the relation between the dependent and independent variables, develop your hypothesis in terms of an implicit model.
2. **Operationalize variables.** Having developed your hypothesis, think of how you can measure the relevant variables. You have to make sure that you are measuring what you intend to measure and nothing else.
3. **Collect clean data.** If there are built-in biases, make sure you are aware of them and can make the necessary adjustments to your model.
4. **Plot the dependent variable against each independent variable.** To formulate the explicit functional form, plot the dependent variable against the independent variables. Determine which form is most appropriate.
5. **Estimate the regression equation** and make necessary adjustments to omit unnecessary variables and include necessary ones.
6. **Check for multicollinearity.** Calculate the correlation matrix. If the problem of multicollinearity is not acute, leave it alone. Otherwise, see if you can pick one variable from the collinear ones to represent the set. Also, see if adding the variables or taking the first difference makes theoretical sense.
7. If there is reason, **check for heteroskedasticity.**
8. **Check for serial correlation.** Use Durbin-Watson statistics, if available.
9. **Present estimation results clearly** and draw conclusions.
10. **Explain all the assumptions** and point out the possible sources of biases in your conclusions.

Tips: Go for parsimony. If two models explain approximately the same amount, choose the one with fewer variables. Moreover, **keep it simple.** Unless a more complicated functional form is truly necessary, choose the simpler one. When forecasting the future, remember that the error of estimation will flare out as you move away from the sample mean. Therefore, the difficulty of making accurate predictions increases exponentially as you go farther into the future.

EXERCISES

1. What is causal prediction? What are its advantages over trend projection? What are its relative shortcomings?
2. Refer to our estimated model of housing price in equation B.3, then comment on the following four units presently on the market:

	Area (square feet)	View	Cul-de-sac	Asking price (\$)
a.	2,500	Yes	No	295,000
b.	3,200	Yes	Yes	325,000
c.	1,500	No	No	162,000
d.	1,950	No	Yes	182,000

To explain housing prices more thoroughly, what are some other variables you would include in the model?

3. Suppose you have been asked to forecast the crime rate of your city. You have decided to use a cross-sectional model of fifty-five cities across the nation.
 - a. What measure of the dependent variable would you use?
 - b. Which variables would you include as the independent variables, and what signs for their coefficients would you postulate?
 - c. What would be the source(s) of your information?
 - d. Explain the various statistical problems that you would face when estimating this model.
 - e. What actions would you take to correct these problems?
4. Suppose your state legislature is considering legalizing a state-run lottery. One aspect of the lottery that is under scrutiny is the demographic profile of the prospective players. A survey of total yearly purchases of lottery tickets by 500 participants in a neighboring state shows the following relation:

$$\begin{aligned} \text{Total purchases} &= 10.51 - 0.086 \text{ PI} + 1.23 \text{ Age} - 0.162 \text{ Ed} + 2.59 \text{ Min} + 3.05 \text{ Male} \\ t &= (6.78) \quad (3.67) \quad (2.46) \quad (2.01) \quad (2.01) \quad (1.14) \\ R^2 &= 0.68 \\ \text{Adjusted } R^2 &= 0.59 \\ F(5, 494) &= 189.95 \\ n &= 500 \end{aligned}$$

where PI is personal income ($\times \$1,000$), age is the age of the lottery player, ed is years of education, min is a dummy variable (minority = 1, nonminority = 0), and male is a dummy variable (male = 1, female = 0).

- a. Write a detailed report explaining the results.
 - b. What are the important policy implications the legislators should be aware of?
 - c. What are some of the other independent variables that could have been included?
 - d. On the basis of this estimated model, how much are you expected to spend on the lottery per month?
5. The water utility department of your town has estimated the following model of water use per capita:

$$\begin{aligned} \text{PWU} &= 15.64 + 1.86 Y + 3.29 \text{ CH} + 2.87 \text{ AD} + 0.029 \text{ SQ} + 0.009 \text{ LS} \\ t &= (2.39) \quad (7.85) \quad (3.99) \quad (1.99) \quad (2.53) \quad (2.01) \\ R^2 &= 0.76 \\ \text{Adjusted } R^2 &= 0.74 \\ F(5, 344) &= 189.95 \\ n &= 350 \end{aligned}$$

where PWU is per capita water use (in gallons), Y is income per household ($\times \$1,000$), CH is the number of children in the house, AD is the number of adults in the house, SQ is the square footage of covered area, and LS is the lot size (land area).

Interpret this equation and write a report explaining the significance of these findings.

- What are the possible sources of bias in the estimation of a classical least squares method? Explain the terms *multicollinearity*, *heteroskedasticity*, and *serial correlation*. Discuss how you would detect them in an estimated relationship. What are some of the ways of eliminating multicollinearity and heteroskedasticity?
- Suppose there are fifteen counties in your state. You are given the current year's sales tax revenue and a number of independent variables. Estimate the equation and, on the basis of the projected growth of the independent variables, forecast the state's total sales tax revenue five years from now. Also, comment on the relative size of the β coefficients. (Hint: Estimate the equation, examine it for various errors, and, on the basis of the correct specification, reestimate the equation. Using the revised equation, forecast each county's tax revenue and then add the county revenue figures to forecast the state's tax revenue.)

County	Current year's data			Forecasted data		
	Tax revenue ($\times \$1$ million)	Population ($\times 1,000$)	Per capita income ($\times \$1,000$)	Land area devoted to agriculture (percent)	Population ($\times 1,000$)	Per capita income ($\times \$1,000$)
A	38.7	50.0	23.2	60.0	52.0	24.0
B	156.0	75.0	32.8	20.0	86.5	36.2
C	115.9	151.1	22.1	75.0	153.0	22.8
D	98.7	45.6	36.9	15.0	54.0	39.7
E	68.3	98.0	24.1	62.0	101.0	24.9
F	220.8	94.0	36.0	18.0	99.0	38.6
G	75.2	102.9	23.0	62.0	105.0	23.2
H	268.8	91.7	41.0	3.0	110.0	46.2
I	32.1	65.0	21.9	82.0	61.0	22.1
J	69.2	47.1	34.6	59.7	49.5	33.6
K	199.9	98.0	35.6	21.0	106.0	37.8
L	86.0	110.0	23.8	58.0	111.8	24.3
M	112.8	202.0	25.0	66.0	198.0	25.0
N	209.0	85.3	40.1	5.0	96.7	46.5
O	116.0	66.0	39.0	22.0	121.0	41.2

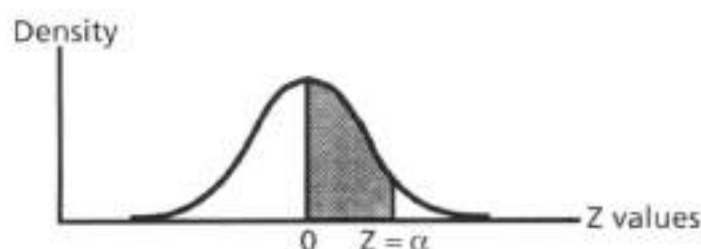
Notes

- For an excellent discussion of computer-assisted mass appraisal of property value, see Glenn W. Fisher, *The Worst Tax? A History of the Property Tax in America* (Lawrence: University of Kansas Press, 1996), 176–186. See also J. Richard Aronson and Eli Schwartz, eds., *Management Policies in Local Government Finance*, 4th ed. (Washington, D.C.: International City/County Management Association, 1996), 211.
- In most states, property tax is 1 percent of the assessed value.
- I did not write the formula for deriving multiple regression coefficients. The formula is complex, and calculating it is extremely time consuming. For estimating multiple correlation,

you should familiarize yourself with available statistical packages. There are a number of excellent and user-friendly software programs, of which SPSS and Minitab are perhaps the most widely used.

4. Our problem has been compounded by the fact that we have a percentage measure as a dependent variable and are therefore restricted between 0 and 1. However, even without this restriction, higher order polynomials can quickly ascend or descend to absurd levels of prediction, particularly for the values of the independent variable outside the sample range.
5. You may test this statement by creating a linearly dependent variable and then including both independent variables in a regression model. Most software will give you an error message stating that the coefficients cannot be computed. A few software programs will give you highly imprecise estimates due to rounding errors.
6. The presence of multicollinearity is best determined by the test of variance inflation factor. However, this test is a bit too complicated for this book. See O. E. Farrar and R. R. Glauber, "Multicollinearity in Regression Analysis: The Problem Revisited," *Review of Economics and Statistics* (1967): 92–107; and David A. Belsley, Edwin Kuh, and Roy E. Welsch, *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity* (New York: John Wiley, 1980). For an excellent overall discussion, see A. H. Studenmund, *Using Econometrics: A Practical Guide*, 2nd ed. (New York: HarperCollins, 1992).
7. This test is based on J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least-Square Regression," *Biometrika* (1951): 159–177.
8. See, for example, Harry H. Kelejian and Wallace E. Oates, *Introduction to Econometrics: Principles and Applications* (New York: Harper & Row, 1981); and Studenmund, *Using Econometrics*.
9. This situation will be close to the simultaneous bias. However, the bias of simultaneity has been deliberately left out of this introductory book. Without explanation, I will simply state that this bias can be corrected by using the instrumental variable approach. If you are interested in learning more about this problem, see Peter Kennedy, *A Guide to Econometrics*, 4th ed. (Cambridge, Mass.: MIT Press, 1998). Also see Studenmund, *Using Econometrics*, or Robert S. Pindyck and Daniel L. Rubinfeld, *Econometric Models and Econometric Forecasts*, 2nd ed. (New York: McGraw-Hill, 1981).

Appendix C

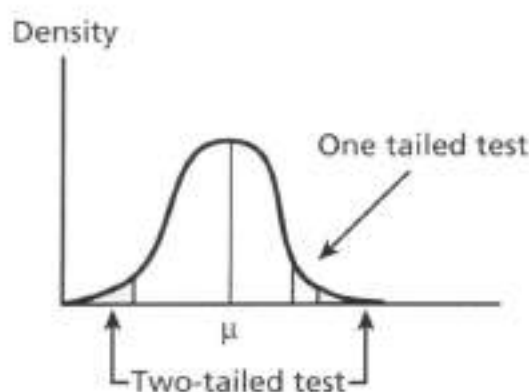


Areas of the Standard Normal Distribution (the Z table)

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2793	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4848	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998

Note: Suppose you want to find the area under the standard normal curve between the values $Z = 0$ and $Z = 1.77$. In that case, read down the first column of Z values and locate the row for $Z = 1.7$. Then read along this row and find the number corresponding to the column 0.07 (in effect, you have added 1.70 to 0.07 to obtain the value for 1.77). This value is .4616.

Appendix D



Critical Values of the t Distribution

Degrees of freedom	Level of significance				
			<i>One-tailed test</i>		
	10%	5%	2.5%	1%	0.5%
	20%	10%	5%	2%	1%
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.898	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
28	1.313	1.701	2.048	2.467	2.763
30	1.310	1.697	2.042	2.457	2.750
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
Infinity (normal distribution)	1.282	1.645	1.960	2.326	2.576

Appendix E

Critical Values of the F Statistic: 5 Percent Level of Significance

		v_1 = degrees of freedom for numerator									
		1	2	3	4	5	6	7	8	10	20
v_2 = degrees of freedom for denominator	1	161	200	216	225	230	234	237	239	242	248
	2	18.5	19.0	19.2	19.2	19.2	19.3	19.4	19.4	19.4	19.4
	3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.79	8.66
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.80
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.56
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	3.87
	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.44
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.15
	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.14	2.94
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.98	2.77
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.85	2.65
	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.75	2.54
	13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.67	2.46
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.60	2.39
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.54	2.33
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.28
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.45	2.23
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.41	2.19
	19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.38	2.16
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.35	2.12
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.32	2.10
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.30	2.07
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.27	2.05
	24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.25	2.03
	25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.24	2.10
	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.16	1.93
	60	4.00	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.08	1.84
	120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.91	1.66
	Infinity	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.83	1.57

Note: The F statistic is a joint probability distribution that measures statistical significance on the basis of two-sided hypotheses about more than one regression coefficient at a time. Unlike the t statistic, the F statistic is measured by two sets of degrees of freedom. The degrees of freedom for the numerator (the column values), v_1 , is calculated by K , the number of restrictions (coefficients for the independent variables plus the intercept term), and the denominator (the row values), $v_2 = n - K - 1$, where n is the number of observations. Thus, if in an estimated equation there are fifty observations and five independent variables, the numerator value (v_1) for the F statistic is $5 + 1 = 6$, and the denominator value (v_2) is $50 - 5 - 1 = 44$. Since we do not have the exact value corresponding to these degrees of freedom, approximate it with the closest number, which is $F(6, 30) = 2.42$.

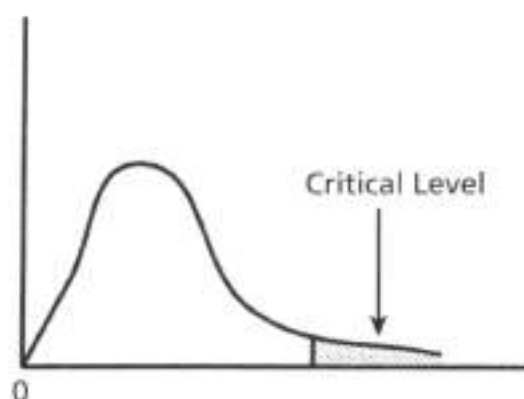
Appendix F

Critical Values of the *F* Statistic: 1 Percent Level of Significance

	v_1 = degrees of freedom for numerator									
	1	2	3	4	5	6	7	8	10	20
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,056	6,209
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.2	26.7
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.5	14.0
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.1	9.55
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.87	7.40
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.62	6.16
8	11.3	8.65	7.59	7.01	6.63	6.37	6.28	6.03	5.81	5.36
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.26	4.81
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.85	4.41
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.30	4.10
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.10	3.86
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	3.94	3.66
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	3.80	3.51
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.69	3.37
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.59	3.26
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.51	3.16
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.43	3.08
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.37	3.00
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.31	2.94
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.25	3.51	2.88
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.26	2.83
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.21	2.78
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.17	2.74
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.13	2.70
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	2.98	2.55
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.80	2.37
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.63	2.20
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.47	2.03
Infinity	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.32	1.88

v_2 = degrees of freedom for denominator

Appendix G



The Chi-square Distribution

Degrees of freedom	Level of significance (probability of a value of at least as large as the table entry)			
	10%	5%	2.5%	1%
1	2.71	3.84	5.20	6.63
2	4.61	5.99	7.38	9.21
3	6.25	7.81	9.35	11.34
4	7.78	9.49	11.14	13.28
5	9.24	11.07	12.83	15.09
6	10.64	12.59	14.45	16.81
7	12.02	14.07	16.01	18.48
8	13.36	15.51	17.53	20.1
9	14.68	16.92	19.02	21.7
10	15.99	18.31	20.5	23.2
11	17.28	19.68	21.9	24.7
12	18.55	21.0	23.3	26.2
13	19.81	22.4	24.7	27.7
14	21.1	23.7	26.1	29.1
15	22.3	25.0	27.5	30.6
16	23.5	26.3	28.8	32.0
17	24.8	27.6	30.2	33.4
18	26.0	28.9	31.5	34.8
19	27.2	30.1	32.9	36.2
20	28.4	31.4	34.2	37.6

Index

Figures and tables are indicated by f and t following page numbers.

- Abandoned Mine Reclamation Fund, 91
- Absolute measure of efficiency, 16
- Abu Ghraib prison abuses, 37
- Academic performance, ranking of nations, 4, 4t
- Academics as analysts, 28
- Accuracy
 - of measuring scales, 178–179, 179f
 - of simple regression analysis, 278–280
- ACLU (American Civil Liberties Union), 84
- Actions, 323
 - alternate courses of, 103–105
- Adjusted R^2 , 303–304
- Adjustment and anchoring as biases, 127
- Adoption lag and delay in policy implementation, 67
- Advanced searches, 194
- Advocacy, 233–234
 - analysts and, 28–34
- Affirmative action plans, 120
- African Americans
 - life insurance of, 223
 - tyranny of the majority and, 21
- Agenda setting models, 76–87
 - cost-benefit analysis, 86–87
 - neglect or deliberate inaction and, 89–91, 90f
 - pluralist vs. elitist models, 83, 83t
 - school safety as example of pluralist model, 84
 - types of, 76–77
- AIDS, information about, 197, 227–228, 230, 234
- Aitkens, Alexander, 317
- Akerlof, George, 67, 68, 73, 74, 343
- Alchain, Armen, 25
- Alexander, Matthew, 44
- Allais, Maurice, 125, 130
- Allais paradox, 124–125, 331
- Allocation, 53
- Alonso, William, 197, 202
- Alternate courses of action, 103–105
 - consistency with goals, 103
 - evaluation of, 104
 - feasibility of, 103
 - forecasting, 103
 - identification of, 103
 - policy goal achievement and, 4–5
- American Civil Liberties Union (ACLU), 84
- American Economic Association, 39
- American Enterprise Institute, 28
- Ames, E., 280, 317
- Analysis of historical data, 249–269, 311–313. *See also* Projection techniques
- Analysts, role and tools of, 27–45
 - client groups of, 28
 - domain of, 29
 - limitations of, 38–39
 - models of analysis, 30–32, 31f
 - quantitative techniques, use of, 35–40
 - statistical techniques used by, 42–43
 - types of analysts, 28–34
- Analytical constraints on role of government, 66–67
- Anchoring as bias, 122, 127
- Anderson, Benedict, 13, 25
- Anderson, James E., 98, 99
- Animal spirit, 67
- Annual Statistical Abstract of the United States*, 196
- Anonymity, 233
- Antilog, 301
- Aronson, J. Richard, 344
- Ascher, William, 45
- Assumption of continuity, 225
- Assumptions regarding human behavior, 38–39
- Asymmetric distributions, 143–144, 143–144t, 143f. *See also* Symmetric distribution
- Asymptotic convergence, 133
- Asymptotic probability, 151
- Audience, knowledge of, 383
- Autoregressive model, 311
- Availability of information, 122, 125–126
- Average price vs. marginal costs, 51

- Axelrod, Robert, 340, 344
 Ayton, Peter, 176
- Baade, Robert A., 129
 Bachrach, Peter, 99
 Bagley, Charles E., 371
 Bailout money, 8, 16
 Baim, Dean V., 129
 Baker, James, 36
 Baldassare, Mark, 344
 Ball, George, 387
 Banking crisis of 2008, 10–11
 Bank insurance, 68
 Banks, Arthur, 202
 Baratz, Morton S., 99
 Bar charts, 144, 210
 Bardach, Eugene, 379, 387
 Bardes, Barbara A., 100
 Bargaining, 341
 Bar graphs, 144, 210
 Barrett, Louise, 25
 Barrett, William P., 202
 Baseball and probability, 152–153
 Base figure, 209
 Battle of the Bismarck
 Sea (game theory case), 335, 335–336f
 Baumol, William J., 25, 74
 Beck, Douglas, 248
 Becker, Gary S., 370
 Behavioral models, 88
 Behn, Robert, 377, 387
 Behrens, William W., III, 269
 Bell curve, 141–142
 Benefits analysis. *See* Cost-benefit analysis
 Bentham, Jeremy, 106
 Bernanke, Ben, 14
 Bernstein, Carl, 36
 Beta coefficient, 276
 Biases
 awareness of, 375–376
 in data/observations, 39
 ideological, 41, 42
 of imaginability, 126
 of reasoning, 122–127, 321
 in statistical interpretation, 214
 systematic, 42
 Biddie, Andrea K., 371
 bin Laden, Osama, 13, 33
 Birkland, Thomas, 83, 85, 99
 Birody, Dan, 130
 "Blackmail" potential, 103–104
 Black Panthers, 9
 Blaug, Mark, 73
 Blimes, Linda, 99
 Boolean searches, 194
 Borger, Mya M., 129
 Bradford, David F., 370
 Brandt, Allan, 44
 Brewer, Gary D., 96, 100
 Brinkley, Joel, 73
 Broad, William, 388
 Brookings Institution, 28, 41
 Brooks, David, 10–11
 Bruce, Neil, 75
 Budget, federal
 deficit as percentage of GDP, 297–299, 297t, 298f, 298t, 299f
 systems vs. incremental approach, 35–36
 2000 presidential election debate over surplus, 266
 Bueno de Mesquita, Bruce, 242, 248, 344
 Buford, Bill, 25
 Bulow, Jeremy, 344
 Burchell, Robert W., 247
 Bureaucracy
 elitist model of agenda setting and, 81, 81f
 pluralist model of agenda setting and, 77, 77f
 Bureau of Labor Statistics, 199, 205
 Bush, George H. W., 36, 243, 245
 Bush, George W.
 Abu Ghraib prison abuses, 37
 antiterrorism actions of, 32
 Enron investigation and, 36–37
 Halliburton and, 118–119
 ideological slant of efficiency and, 17–18
 neocon counterterrorism policy of, 33
 9/11 terrorist attacks, 22
 presidential debates (2000), 266
 presidential election (2004), 12
 Butts, R. E., 344
- Cagan, Joanna, 129
 California. *See also* San Diego, California
 defense expenditures, 120
 DNA testing of criminal defendants, 91
 electricity deregulation, 68
 gold rush, 292
 immigration controversy, 51
 Mammoth, and predictions of volcano eruption, 265
 milk producers, 56
 Orange County bankruptcy, 332–333
 Proposition 13 (tax revolt), 70, 332–333
 Campaign contributions, 282–283
 Capital punishment. *See* Death penalty
 Carter, Douglas, 98
 Carter, Jimmy, 12, 91, 121
 Case study method, 230
 Casti, John, 38, 45, 99, 176, 344
 Catastrophic trend, 255–256, 256f
 Categorical precedence and pluralist model of agenda setting, 80
 Cato Institute, 28
 Causal connections/links, identification of, 109, 110f, 114, 225

- Causal regression model, 353–354
- CDC (Centers for Disease Control), 78, 197
- Census Bureau, 40
- Census Bureau Catalog and Guide*, 196
- Census of Governments*, 196
- Center on Budget and Policy Priorities, 267
- Centers for Disease Control (CDC), 78, 197
- Central Intelligence Agency (CIA), 21, 197, 245
- Central limit theorem, 155–156
- Central tendency, 131, 132–137, 144
- Chances, failure to self-correct and dilution, 123–124
- Chandler, Ralph C., 130
- Changes in slope, 292–296, 292*t*, 293*f*, 295*f*, 296*t*
- Changes in trend, 288–292, 288*t*, 289–290*f*
- Chatterjee, Pratap, 129
- Cheney, Dick, 18, 33, 37, 118–119
- Cheung, Richard, 344
- Chi-squared test, 165–169
- Choi, Young Back, 321, 343
- CIA. *See* Central Intelligence Agency
- Gialdini, Robert, 26
- Citizen activist groups
- elitist model of agenda setting and, 81, 81*f*
 - pluralist model of agenda setting and, 77, 77*f*
- Citron, Robert, 332–333
- Clark, Christie, 75
- Clarke, Richard, 33
- Class divisions in presidential voting, 132*f*
- Classification of goods, 6–7, 6*t*
- Clayton Act of 1914, 54
- Clean Air Act of 1990, 37, 71
- Clines, Francis X., 99
- Clinton, Bill
- banking regulations and, 11
 - navy ships, 64
 - oil industry and, 55
 - terrorist attacks and, 33
- Club goods (common pooled resources), 6*t*, 7
- Cluster samples, 186
- Coalitions formed around institutional agendas, 92, 93*f*
- Coase, Ronald, 71
- Cobb, Roger, 77, 80, 98
- COE (U.S. Army Corps of Engineers), 325, 343
- Coefficient for trend (slope), 275
- Coefficient of determination. *See* R^2
- Coefficient of variation, 141
- Coefficients
- beta coefficient, 276
 - relevance of estimated coefficients, 280–283
 - significance of individual coefficients, 283–286, 286*t*
 - significance when taken together, 304–305
- Coequal (nonhierarchical) model, 31, 31*f*
- Cohen, David K., 21, 36–37, 45
- Cohen, Richard, 99
- Collapse* (Diamond), 9
- Collection bias, 182
- Collection costs, outweighing benefit of charging fees, 72
- Collection of data. *See* Sources of data; Surveys
- Collective action problem, 8
- Collectively exhaustive inclusion of alternatives in decision tree, 324
- Columbine High School incident, 84
- Column marginal, 166
- Common good, 39
- Common log, 301
- Common pooled resources (club goods), 6*t*, 7
- Community security, 23
- Comparable city method, 230
- Comparison of different series of numbers, 162–164
- Compound interest, 355
- Conditional forecasting, 306–307
- Confidence interval, 159–161
- Conflict, 89–91
- Conflict in values, responses to, 32–34, 34*f*, 36–37
- Conflict of interest, 42
- Congress
- agenda setting by, 77
 - policy adoption considerations, 91–92
 - self-interest and ideology as influencing, 92
- Congressional Budget Office, 198, 266, 362, 375
- Consequences, forecasting, 116–117
- Consistency with goals, 103
- Consolidated policy change, 95
- Constant, 133
- Constant vs. current dollars, 204–208, 204*t*, 205*f*, 207*f*, 208*t*
- Constitution, U.S., 5
- Fourth Amendment, 37
- Consumer price index, 206–207, 206*t*, 214
- Consumer surplus, 50, 50*f*, 346, 350, 350*f*
- Consumption efficiency, 70
- Contingency table, 165, 166*t*
- Control variables, 89, 109, 110*f*
- Convincing decision makers, 385–386
- Core values of policy coalitions, 92
- Correlation analyses, 114, 114*t*
- Correlation coefficient (Pearson's r), 145–147, 146*f*, 147*t*, 167

- Corrigan, Tim, 224
- Cost-benefit analysis, 345–371
 agenda setting using, 86–87
 case study using, 367–368
 choice of best alternative, 364–368, 365t
 choice of discount rate, 360–362
 classifications of, 348, 348f
 defining goals, 346
 estimation of costs and benefits, 349t, 350–355
 examples of, 120
 forecasting using, 88–89
 identifying alternatives, 346–347
 intangible costs and benefits, 348, 351–353, 354
 internal rate of return, 362–363, 363f
 listing alternatives' costs and benefits, 347–348
 measurement of future loss or gain, 353–355
 policy cycle and, 5
 present value analysis, 355–363. *See also* Present value
 public goods allocation and, 7
 ratio, 364–365
 redistribution of income, 366
 social vs. private, 346–347
 steps in process of, 346
 valuation of costs and benefits, 349t, 350–355
- Cost-effectiveness analysis, 366–368
- County and City Data Book*, 196
- County Yearbook*, 196
- Court agenda setting, 77
- Credible commitment problem, 8
- Crime rates, 114
 autoregressive model to forecast, 311–312, 312t
- Critical thinking and research design, 101–130
 challenges to critical thinking, 122–127
 choosing method of analysis, 109–115
 designing policy research, 117–122
 goal setting, 104–109
 objective analysis, 102–104
- Critical value, 158
- Cross-section analyses, 114
- Cullen, Dave, 99
- Cultural constraints on role of government, 65
- Cumulative probability distribution, 151, 153t
- Current vs. constant dollars, 204–208, 204t, 205f, 207f, 208t
- Curve fitting, forecasting by, 312–313
- Cyclical trend, 252
- Damasio, Antonio, 11–12, 25
- D.A.R.E. (Drug Abuse Resistance Education), 88
- Databases, 195
- Data sources, 177–202. *See also* Sources of data
- Davis, Robert K., 130
- Death penalty, 90–91, 95
- Deceptive use of statistics, 213–222
- Decision making
 choice of best alternative, 364–368, 365t
 criteria for choosing among alternatives, 104
 expected payoff as factor, 321–322, 330–332, 331f, 334t
 game theory, 332–339. *See also* Game theory
 graphical methods in, 211–212
 rules to apply under uncertainty, 338–339
- Decision tree analysis, 322–332
 actions, 323
 chance nodes, 323, 325
 collectively exhaustive inclusion of alternatives, 324
 decision nodes, 323, 325
 do-nothing option, 322–323, 325, 327
 folding backwards, 330
 mutually exclusive alternatives, 324
 outcomes, 322–323
 payoffs, 323, 327, 330–332, 331f, 334t
 probabilities, 323, 327
 structuring of tree, 324–325
- Declaration of Independence (U.S.), 18, 63
- Dede, Martha J., 248
- de Graaf, Janny, 84, 99
- Degrees of freedom, 140, 161
- DeLeon, Peter, 37, 44, 45, 96, 100
- Delphi technique, 232–237
- deMause, Neil, 129
- Democracy, limits of, 63–64
- Demsetz, Harold, 25
- Dependent variables, 109, 110f, 225, 275
 lag dependent variables, 311–313
- Deregulation, 68–69
- Descartes, René, 3
- Designing policy research, 117–122
- Diamond, Jared, 9
- Dickey, Christopher, 26
- Difference of proportions test, 164–165
- Diffuse costs and benefits, 86–87, 86t
- Direct costs and benefits, 348
- Direct infusion of money, 69
- Direct observation to collect primary data, 179
 errors due to, 182–184

- Disasters' effect on agenda setting, 83–84
- Disclosure of information, 57
- Discounted future earnings method, 352
- Discount rate, 356–359
choice of, 360–362
plot of present values as function of, 359f
time preference and, 357f
- Discrimination, 120, 223
- Dispersion
of distributions, 141–142, 141t
ex post data, analysis of, 131
measures of, 137–144, 152
normal distributions with various levels of, 155f
- Disraeli, Benjamin, 213
- Dissertation Abstracts/Digital Dissertations*, 200
- Distribution
bell-shaped distribution, 141–142
dispersion of distributions, 141–142, 141t
as government role, 53
joint probability distribution, 167
probability distribution, 151–157, 153t, 154f
skewed distributions, 141–144
of statistical summary, 233
stem-leaf method, 144–145, 145f, 145t
Z distribution, 158
- Dixit, Avinash, 340–341, 343
- DNA testing to determine guilt, 90–91, 96
- Dominant strategy, 337, 338–339
- Do-nothing option, 322–323, 325, 327
- Douglas, Susan, 26
- Downs, Anthony, 89, 90, 99
- Druckman, James, 25
- Drug Abuse Resistance Education (D.A.R.E.), 88
- Drugs, illegal
education about, 88, 115–116, 211–212, 212f
forecasting use of, 115–116
long-term programs for dealing with, 359
regulation, 87
trafficking, 56, 58, 87, 359
- Drunk driving, 87
- Dubnick, Melvin J., 100
- Dummy variable, 288–292, 290t
- Dunbar, Robin, 25
- Dunn, William, 234, 248
- Dye, Richard F., 129
- Dye, Thomas R., 81, 98
- e* (equal to 2.71828), 301
- Easton, David, 45
- Ebbert, Stephanie, 73
- EBSCOhost Academic Search Elite, 196
- Economic analyses, fundamentals of, 49–52
- Economic cycles, 252, 252f
- Economic rationality, 50–51
- Economic security, 23
- Education. *See* Schools
- Effectiveness of presentation, 382–386
- Effectiveness of search, 126
- Efficiency as policy goal, 15, 15f, 16–18
- Efficiency-equity trade-off, 17, 19
- Elder, Charles D., 98
- Elections
limits of democracy, 63–64
polls, 182
predictions, 243
2000 presidential election, 216–217, 266
2004 presidential election, 12
- Electricity deregulation, 68
- Elitist model of agenda setting, 77, 81–83, 81f
- Elkind, Peter, 44
- Ellerman, A. Denny, 75
- Ellison, Brian, 93, 100
- Eminent domain, 120
- Emissions trading, 71
- Endangered Species Act, 120, 121
- Enron investigation, 36–37
- Entrepreneurs, political, 13
- Environment
Animas-La Plata River project, 93
Clean Air Act of 1990, 37, 71
endangered species, 120, 121, 353
energy and dam construction, 121
Exxon Valdez settlement, 353
knowledge constraints on role of government, 66
PCBs, 352–353
- Environmental security, 23
- Equilibrium, 48–49, 48f
- Equity as policy goal, 15, 15f, 17, 18–19
- Equity-efficiency trade-off, 17, 19
- Equity in distribution of income, 19
- Erickson, Carrolly, 44
- Errors. *See* Biases; Estimation errors; Random errors; Systematic errors in sampling
- Error term, 277
- Estimation errors, 274, 274f, 276
- Estimation of costs and benefits, 349t, 350–355
- Ethics
public policy analysis and, 9–14, 40–42
value conflict and, 32–34, 34f, 36–37
- Evaluation
of policy, 93–94
of possible outcomes, 104

- Evidence chosen to prove event, 124
- Ex ante facto, 131
- Ex ante forecasting, 306
- Excludability as classification of goods, 6, 6*t*
- Exhaustibility (rivalrous quality) of a good, 6, 6*t*
- Expected payoff and decision making, 321–322, 330–332, 331*f*, 334*t*
- Expected utility model, 242–244, 243–244*f*
- Experience goods, 59
- Experimental research designs, 110–112. *See also* Quasi-experimental research designs
- Exponential trend, 254–255, 255*f*
- Exports and imports institutional constraints on, 65–66 mercantilism and, 47
- Ex post facto, 131
- Ex post forecasting, 306
- Externalities awareness of, 376–377 negative externalities, 59–61, 71, 104, 347 positive externalities, 59–61, 103, 347, 362 social costs and, 59–61
- External threats to security, 21
- Exxon Valdez settlement, 353
- Face-to-face interviews, 189
- Face value of life insurance method, 352
- Factor mobility, 55
- Factual information, 88
- Fairness, 121–122
- Falsifiability, 5
- Family Support Act of 1988, 37
- FAT. *See* Feasibility assessment technique
- FDA (Food and Drug Administration), 76, 168
- FDIC (Federal Deposit Insurance Corp.), 10, 68
- Fear, 339–340
- Feasibility assessment technique (FAT), 238–241 calculated influence, 241*t* numbers, origin of, 244 potential for influence, 240*t* relative issue positions, 238*t*, 239*f* resource availability, 239*t* resource rankings, 240*t*
- Feasibility of alternatives, 103
- Federal budget. *See* Budget, federal
- Federal debt, 29
- Federal Deposit Insurance Corp. (FDIC), 10, 68
- Federal Emergency Management Agency (FEMA), 325
- Federal Reserve, 14, 37
- Feedback, 94
- Feingold, Stephen, 197, 202
- Feldman, Martha, 388
- Felt, Mark ("Deep Throat"), 36
- FEMA (Federal Emergency Management Agency), 325
- FIA (fiscal impact analysis), 229–230, 229*t*
- Financing of sports stadiums, 101, 103–106
- Fineberg, Harvey V., 343
- Fiscal impact analysis (FIA), 229–230, 229*t*
- Fischer, Frank, 45
- Fisher, Anthony C., 130
- Fisher, Roger, 341, 344
- Fisher, Ronald A., 317
- Fixed weight, 205
- Flood damage, 325–330 decision tree for flood control, 326*f*, 328*f* decision tree with total expected costs, 329*f* estimation of, 325*t*, 327*t* probability of, 326*f* total estimated cost, 330*t*
- Florida Palm Beach County's growth management policies, 191 2000 presidential election results, 216–217
- Focus events, 85
- Focus groups, 190–192
- Food and Drug Administration (FDA), 76, 168
- Food security, 23
- Forbes magazine, 29
- Ford, Gerald, 91, 319–320, 322–324
- Forecasting, 88–89. *See also* Probability; Projection techniques alternate courses of action, 103 by curve fitting, 312–313 history, inadequacy of, 225–226 interval forecasts, 307–311, 310*f* least squares, 271–318. *See also* Least squares method multiple regression, 272, 287–296. *See also* Multiple regression point forecasts, 306–307 politics of, 265–267 problems of, 305–311 random errors in, 273–274, 274*f* without relevant information, 124
- Foucault, Michel, 24
- Fourth Amendment, 37
- Fowler, James, 26
- Framers of Constitution, 30
- Framing, 13
- Frantz, Roger S., 73, 176, 387
- Freed, Stanley, 129
- Freeman, J. L., 98
- Free market correction by government, 53

- government and, 47–74
 market failure, 52–62
 perfectly competitive market, 53
 Free riders, 8, 9, 87
F test, 305
 Functional termination of policy, 96
 Fundamental issue, identification of, 103

 Galton, Francis, 124
 Game theory, 332–339
 defection and its consequences, 339–340
 dominant strategy, 337, 338–339
 local government decision making using, 337–338
 minimax strategy, 334, 337
 Nash equilibrium, 335, 337
 prisoner's dilemma, 337–338, 338*t*
 tit-for-tat strategy, 340
 zero-sum games, 338–339
 Gandhi, 13
 GAO (Government Accountability Office), 196, 362
 Garbage can model of agenda setting, 82–83
 Gart, Alan, 75
 Gauss, Carl Friedrich, 317
 GDP
 deflator, 206
 federal budget deficit as percentage of, 297–299, 297*t*, 298*f*, 298*t*, 299*f*
 per capita, 271–273, 272*f*, 288, 304
 Geanakoplos, John, 344
The General Theory of Employment, Interest and Money (Keynes), 43
 General will, 105
 Geneva Convention, 22, 37
 Gillespie, Kate, 73
 Glaberson, William, 100
 Glaeser, Bernhard, 129
 Glass-Steagall Act of 1933, 11
 Gleick, James, 179, 202
 Glezer, Vadim D., 248
 Goals
 consistency with, 103
 for cost-benefit analysis, 347
 of government intervention, 15–23, 15*f*
 setting, 104–109
 Goldhamer, Herbert, 44
 Goleman, Daniel, 25
 Gonzales, Elian, 78–79, 80
 Gore, Al, 266
 Gossett, W.S. (*Student*), 161
 Gould, Stephen Jay, 224
 Government Accountability Office (GAO), 40, 196, 362
 Governmental agendas, 76–77
 Government decision making and game theory, 337–338
 Government-funded institutes, 28
 Government publications, 197
 Government rate of return, 362
 Government role
 failure of government, 62–67
 in free market, 47–74
 goals desired in policies of, 15–23, 15*f*
 market correction, 53
 pervasiveness in daily life, 1–2
 Government staff analysts, 28
 Gramlich, Edward, 370, 371
 Graphical methods, 203–212
 choosing type of, 210–211, 210*t*, 211*f*
 current vs. constant dollars, 204–208, 204*t*, 205*f*, 207*f*, 208*t*
 decision making and, 211–212
 index, 209–210, 210*f*
 percentage change, 208–209, 208*t*
 pictorial presentation of data, 218–219*f*, 218–220, 220*t*
 tabular presentation of data, 221–222, 222*t*
 Gray, Scott, 176
 Great Depression, 3, 10, 32, 38, 52
 Great Leap Forward plan, 17
 Great Society program, 62, 292
 Greed, 339–340
 Greek characters for statistical terminology, 140, 277
 Grossbard-Shechtman, Shoshana, 370
 Gross domestic product. *See* GDP
 Group consensus, 233
 Grouped data, 136–137, 137*t*
 Group identity, 12–14, 183
 Groupthink, 232
 Growth, 53
 Guantanamo Bay detention center, 22, 32
 Gun control. *See* Violence
 Gupta, Dipak K., 25, 26, 99, 202, 343, 387
 Gurr, Ted, 202

 Hagen, Ole, 130
 Haidt, Jonathan, 14, 26
 Halliburton, 118–119
 Hamm, Mark, 26
Handbook of Labor Statistics, 196
 Hanrahan, Jennifer, 98, 100
 Haq, Mahbub ul, 23
 Hardin, Garrett, 25
 Harsanyi, John C., 344
 Haskins, Ron, 41
 Haveman, Robert H., 370
 Hawthorne experiment, 183
 Health care costs, 120, 225–226, 226*f*
 Health security, 23
 Heath, Chip, 388
 Heath, Dan, 388
 Heclo, H., 98

- Heineman, Robert A., 44
 Heisenberg, Werner, 182, 382
 Held, Virginia, 45
 Helms, Jesse, 197
 Henry, Clement M., 73
 Henry, Patrick, 20
 Heritage Foundation, 40
 Heuristics, 122
 Hicks, John, 107
 Higher-order polynomials, 299–301, 300f
 Highway construction, 366
 Hinges, in statistics, 137, 139f
 Hintikka, J., 344
 Hirschman, Albert, 32, 34f, 44
 Histograms, 144
 Historical data analysis, 249–269, 311–313. *See also* Projection techniques
 Hoffman, Bruce, 202
 Hoffman, Frederick L., 223, 224
 Holmstrom, Bengt, 25
 Holtz, Robert Lee, 269
Homo economicus (self-utility-maximizing economic being), 8, 9, 13, 23, 39, 67
 Honest graft situations, 42
 Hoover Institution, 28
 Horizontal trend, 251, 251f
 Housing prices, 133t, 134
 Hudson, Michael, 202
 Huff, Darrell, 215, 218, 221, 224
 Human behavior
 behavioral models, 88
 economic rationality to predict, 50
 Human security, factors of, 23
 Hurricane Katrina, 21, 83
 Hussein, Saddam, 33, 245
 Hypothesis testing, 157–165
 confidence interval and, 159–161
 critical value, 158
 level of significance, 157
 one-tailed test, 158, 159f
 quantitative techniques of analysis and, 38
 risks in, 167–169, 168t
 steps in, 168–169
 t test, 161–162
 two-tailed test, 160–161, 161f, 162
 ICMA (International City/County Management Association), 182
 Ideological bias, 41, 42
 Ideologies, 28–29
 Illegal immigration, 182, 231
 Illusory correlation, 126
 Imaginability, 126
 Implementation of policy, 5, 67, 92–93
 Inadequacy of knowledge and uncertainty, 321
 Income
 nominal income, 206
 real income, 206
 as welfare proxy, 118–119
 Independent probability, 150
 Independent variables
 intercept dummy variable, 288–292, 290t, 292
 irrelevant, 303–304
 method of analysis, 109, 110f
 prediction, 225
 regression model and, 275
 Index, 209–210, 210f
Index to Current Urban Documents, 200
 Indifference map, 357
 Indirect costs and benefits, 348
 Indivisibility, 56
 Inefficiency, 18
 Inflation, 206–207
 Information disclosure restrictions, 57–59
 In-house political analysts (staff analysts), 30–32
 Input-output ratio (efficiency), 16
 Institute for Social Research at University of Michigan, 198
 Institutional agendas, 76–77
 conflict's effect on, 89–91
 nondecision's effect on, 89–91, 90f
 pluralist model and, 77–80, 77f
 Institutional constraints on role of government, 65–66
 Insurance
 African Americans and life insurance, 223
 bank insurance, 68
 face value of life insurance method, 352
 Intangible costs and benefits, 348, 354
 Integrity, 41–42
 Intercept dummy variable, 288–292, 290t
 Intercept term, 275, 285
 Interdependence of human actions and uncertainty, 321
 Internal rate of return, 362–363, 363f
 Internal security, 22
 International City/County Management Association (ICMA), 182
 International Labour Organization, 28
 International Monetary Fund, 28
 Internet searches, 193–200, 194f
 Interval forecasts, 307–311, 310f
 Intervals, 136–137
 Interval scales, 178
 Intervening variables and research design, 112
 Interviews. *See* Opinion polls
 Inverse forms, 302–303, 303f
 Invisible hand of market forces, 3, 21, 48–49
 Iron triangle, 82
 Irrelevant information, 123
 Iteration, 233

- Jacobson, Louis, 100
 Janis, Irving, 248, 376, 387
 Jenkins-Smith, Hank, 92, 98, 99, 100
 Jensen, Ronald, 26
 Jim Crow legislation, 21
 Jodice, David, 202
 Jodice, Michael, 182
 Johnson, Arthur, 129
 Johnson, Kirk, 370
 Johnson, Lyndon B., 35, 62
 Joint probability distribution, 167
 JSTOR, 195, 196
 Judgmental methods
 problems of, 245
 of projection, 227, 232–245
 of sampling, 185
 Judicial impartiality, 282–283
 Kahneman, Daniel, 14, 26, 122, 124, 125, 130
 Kaldor, Nicholas, 107, 129
 Kaldor-Hicks compensation
 principle, 107–108, 119
 Kaletsky, Anatole, 39
 Kalt, Joseph, 91, 99
 Kant, Immanuel, 3
 Kantrow, Buster, 75
 Kau, J. B., 99
 Keefer, Philip, 26
 Kelly, Raymond, 21
 Kemperer, Paul, 344
 Kennedy, Peter, 317
 Kenney, George C., 335–337
 Kernell, Samuel, 26
 Kerry, John, 12
 Keynes, John Maynard
 animal spirit, 67
 government role in market
 correction, 3, 52, 69
 ideas, power of, 43, 45
 policy termination, 96
 political constraints, 65
 Keynesian economics, 32
 Keyword search, 194
 Khemani, Stuti, 26
 King, A., 98
 King, Martin Luther, Jr., 9, 13
 Kingdon, John, 82–83, 98
 Kinsey, Alfred, 387
 Klahr, David, 248
 Klein, Joe, 20
 Knight, Frank, 320, 343
 Knowledge
 of audience, 383
 constraints on role of
 government, 66
 inadequacy of, and
 uncertainty, 321
 ordinary knowledge as
 approach to setting
 policy, 374–375,
 383–384
 Kotovsky, Kenneth, 248
 Kranton, Rachel E., 343
 Kreps, David M., 344
 Krutilla, John V., 130
 Kuran, Timur, 248, 343
 Labor unions and game theory,
 334–335, 337
 Lad, Frank, 169, 176
 Lag dependent variables,
 311–313
 Lagging, 261, 263
 Lakoff, George, 5, 24
 Lane, Robert, 122
 Laplace, Pierre-Simon de, 317
 Lasswell, Harold, 35, 44
 LaVelle, Phillip J., 129
 Law enforcement, 92. *See also*
 Violence
 Law of large numbers, 155, 184,
 281, 282, 284, 286
 Lay, Kenneth, 36
 Least squares method, 271–318
 logic of, 272–275
 multiple regression analysis,
 272, 287–296. *See also*
 Multiple regression
 nonlinear trends, 296–305
 simple regression analysis,
 271, 275–287, 277–
 278*t*. *See also* Simple
 regression model
 Legal constraints on role of
 government, 66
 Legalization, 68
 Legalization of marijuana, 20
 Legendre, Adrien-Marie, 317
 Leibenstein, Harvey, 54, 73
 Lerner, Daniel, 44
 Lester, James P., 98, 100
 Level of significance, 157
 Lewis, Michael, 176
 Lexis/Nexis Academic
 Universe, 196
 Liberty as policy goal, 15, 15*f*,
 17, 20–21
 Liberty-security trade-off,
 17, 22–23
 Lichbach, Mark I., 45
 Liebenstein, Harvey, 343
 Lieberman, Trudy, 98
 Life insurance
 African Americans and, 223
 face value of life insurance
 method, 352
 Likert scales, 190
 Lilla, Mark T., 43, 45
 Lincoln, Abraham, 13
 Lindblom, Charles, 36–37, 45
 Linear functional form
 choice of, 305
 simple regression model of,
 275–287, 277–278*t*
 Linear policy change, 95
 Linear trend, 250, 251*f*
 Liquidity, 332
 Listokin, David, 247
 Literature review. *See* Sources
 of data
 Liu, Zinan, 75
 ln (natural log), 301
 Lobbyists, 82, 82*f*
 Local government decision
 making and game theory,
 337–338
 Logarithms and logs,
 301–302, 302*f*
 Logging moratorium, 353
 Logistic trend, 254–255, 255*f*
 Log-log functions, 301

- Log-transformed forms, 301–302
- Lotteries, 121
- Lower hinge, 137
- Lycett, John, 25
- Lying in response to survey questions, 182
- Lynd, Robert, 35, 44
- Lynk, E. L., 75
- Mabey, Nick, 98
- MacArthur, Douglas, 335
- Mafia, 340
- Mailed questionnaires, 189
- Mainzer, Lewis C., 44
- Majone, Giandomenico, 130
- Makridakis, Spyros, 269
- March, James G., 81, 98
- Marcos, Ferdinand, 243–244
- Mares, David, 73
- Marginal, 166
- Marginal analysis, 49–51, 50*f*
- Marginal costs and benefits of public goods, 64–65
- Marginal revenue, 50–51
vs. average price, 51
- Margolis, Howard, 99, 343
- Margolis, Julius, 370
- Mariani, Mack, 148
- Marijuana legalization and commercialization, 20
- Market. *See* Free market
- Market cost, 60
- Market failure, 52–62
barriers to entry and exit, 55–57
externalities and social costs, 59–61
lack of competition, 53–55
restricted flow of information, 57–59
rising service costs and, 61–62
- Markov, Andrey, 317
- Martin, Clyde, 387
- Massachusetts, housing crisis in Boston, 47, 52
- Master's theses as sources of data, 200
- Mayer, Jane, 26, 44
- Mayo, Elton, 183
- Mazmanian, D., 99
- McCafferty, Stephen, 45
- McCall, Madhavi, 282–283, 318
- McCollum, Kelly, 75
- McDermott, Rose, 26
- McGee, Victor, 269
- McLean, Bethany, 44
- McNamara, Robert, 388
- Meadows, Dennis L., 269
- Meadows, Donella, 269
- Mean, 133–134. *See also* Central tendency
comparison, 164, 164*t*
population mean, 281
of probability distribution, 152, 154*f*
projecting by the mean, 263
skewed distributions and, 144, 144*t*
- Mean absolute deviation, 139, 139*t*
- Median, 134. *See also* Central tendency
skewed distributions and, 144, 144*t*
stem-leaf method and, 144
- Medicaid, 17
- Meltsner, Arnold, 44
- Mercantilism, 47
- Michaels, Meredith, 26
- Microeconomics, 51
- Microsoft antitrust case, 53–54
- Midsread, in statistics, 137, 139*f*
- Mikesell, John, 371
- Milgram, Stanley, 183, 202
- Military
expenditures, global, 21, 22*f*
external factors in valuing loss of life, 352–353
use of game theory, 335–337
- Milk producers in California, 56
- Mill, John Stuart, 20–21, 42
- Miller, Gary, 25
- Minimax strategy, 334, 337
- Minimum wage, 56
- Mitchell, E. J., 99
- Mixture-in-administration model, 31–32, 31*f*
- Mixture-in-policy model, 31*f*, 32
- Mode, 134. *See also* Central tendency
- Models of analysis for public policy
choice of right model, 115–116
coequal (nonhierarchical) model, 31, 31*f*
definitions of, 115
examples of, 31*f*
limitations of, 38–39
mixture-in-administration model, 31–32, 31*f*
mixture-in-policy model, 31*f*, 32
policy-administration dichotomy model, 30, 31*f*
scientific analysis, steps in, 5 types, 30–32
- Monopolies, 53–55
- Monroe, Kristen Renwick, 25
- Moral hazard, 8, 10–11
- Moral Majority, 241
- Morgenstern, Oscar, 321
- MOSAIC-2000, 84
- Moving average, 260, 263–264, 264*f*
- Mueller, John, 99
- Multimodal response, 234
- Multinational agencies, analysts and, 28
- Multiple regression, 287–296.
See also Estimation errors
beta coefficient, 276
changes in slope, 292–296, 292*t*, 293*f*, 295*f*, 296*t*

- changes in trend, 288–292,
288*t*, 289–290*f*
defined, 272
Municipal Yearbook, 196
Mutually exclusive options in
decision tree, 324
Mutually exclusive probability,
150
Myths, 126
- Nachmias, Chava, 129
Nachmias, David, 129
Naive projection, 261–263, 262*t*
Nalebuff, Barry, 340–341, 343,
344
Nash, John, 335
Nash equilibrium, 335, 337
National Center for Public
Policy and Higher
Education, 60, 61
National Institutes of Health
(NIH), 28
National Public Radio (NPR),
8, 14
National Science Foundation,
28
Natural log, 301
Needs assessments, 180
Negative correlation, 145–147
Negative externalities, 59–61,
71, 104, 347
Negative incentives, 88
Negatively skewed distribution,
142, 144
Negative trend, 249
Nelson, Barbara, 98
Net social benefit, 351
Neustadt, Richard E., 343
Newman, David, 248
“New public administration” of
1960s, 30–31
New York City and labor
unions, 337
New York Police Department
(NYPD), 21
New York Times, 10–11
NIMBY (Not In My Backyard)
factor, 14, 353
9/11 terrorist attacks, 21, 22, 33
Nixon, Richard
illegalities by, 36
job performance of, 12
Watergate crisis, 126
No-bid contracts, 18
Nominal income, 206
Nominal scales, 177
Nondecision, 89–91, 90*f*
Nongovernmental
organizations, 28
Noninstitutional agendas,
76–77
Nonlinear policy change, 95
Nonlinear trends, 296–305
inverse forms, 302–303, 303*f*
irrelevant independent
variables, 303–304
log-transformed forms,
301–302, 302*f*
polynomial forms, 296–299
Non-parametric tests, 167
Nonrepresentative samples, 181
Normal distribution, 152, 155*f*,
281, 281*f*
Null hypothesis
coefficients and, 284, 305
defined, 156–157
rejection of, 159, 162, 285,
305
- Obama, Barack
antitrust policy, 55
banking crisis and, 10–11
economic stimulus, 3, 69
equity-efficiency trade-off
and, 19
group identity and, 13
job performance of, 12
stimulus money and, 29
Objective analysis, 42–43
Objective probability, 149–169
Objective professionalism, 102
Objective sampling, 185–186
Observation, number needed
for simple regression,
286–287, 287*f*
Observed frequency, 149–150
- Office of Management and
Budget (OMB), 39, 221,
362
Ohio school violence and
safety, 84
Oil industry, 55
Oligopoly, 55, 56
Olsen, Johan P., 81, 98
Olson, Mancur, 8, 9–10, 25,
87, 99
One-shot case studies, 114
One-tailed test, 158, 159*f*
“On Liberty” (Mill), 20
Online searches,
193–200, 194*f*
Open society and problems of
formulating policy for,
113, 114
Opinion polls, 180. *See also*
Sources of data
methods, 189
survey instruments, 187–189
Opportunity costs, 16, 49, 52,
360
Ordinal scales, 178
Organizational termination of
policy, 96
Ostrom, Elinor, 24, 39, 45
Ostrowski, John W., 248
Outcomes
decision tree analysis,
322–323
evaluation of possible
outcomes, 104
forecasting, 116–117
policy outcomes, 94
target outcome, 150
Outsourcing, 68, 69
- Pabushka, Alvin, 248
PAIS. *See* Public Affairs
Information Service
Pandemics, 78–79
Pareto, Vilfredo, 106
Pareto inferior, 107
Pareto principle, 106–107,
107*f*, 119
Pareto superior, 107

- Passage of time, effect on
 research design, 114–115
 Patton, Carl V., 129
 PCBs, 352–353
 Pearson, Clifford A., 202
 Pearson's r , 145–147, 167
 Pecuniary effect, 351
 Pendleton Act of 1883, 30
 Per capita multiplier method,
 230
 Percentage change, 208–209,
 208*t*
 Perfectly competitive
 market, 53
 Performance evaluation, 94
 Personal security, 23
 Persuasion
 of decision makers, 385–386
 by statistics, 384–385
 Pew Research Center, 12, 22
 Ph.D. dissertations as sources
 of data, 200
 Pictorial presentation of
 statistics, 218–219*f*,
 218–220, 220*t*
 Pie charts, 210
 Pill, Juri, 248
 Pindyck, Robert S., 176
 Planning-Programming
 Budgeting System (PPBS),
 35, 62
 Plano, Jack, 130
 Plato, 3
 Pluralist model of agenda
 setting, 77–80, 77*f*, 85
 Point forecasts, 306–307
 Policy-administration
 dichotomy model, 30, 31*f*
 Policy adoption, 67, 91–92
 Policy analyst's role, 373–374,
 374*t*
 Policy change, 94–95
 Policy cycle, 3–5, 75, 76, 76*f*
 Policy Delphi technique,
 233–234
 Policy evaluation, 93–94
 Policy formation, 75, 76, 76*f*,
 87–91, 377–382
 Policy objectives, 109, 110*f*
 Policy outcomes, 94
 Policy output, 94
 Policy termination, 96–97
 Political constraints on role of
 government, 65
 Political entrepreneurs, 13
Political Science and
 Government, 200
 Political security, 23
 Politics
 defined, 2
 rationality and, 2–5
 Polling methods, 189
 Pollution. *See* Environment
 Polynomial forms, 296–299
 Pomeroy, Wardell, 387
 Pontell, Henry N., 344
 Popper, Karl R., 5, 40–41,
 45, 129
 Population, 133, 180
 Population mean, 281
 Population standard
 deviation, 281
 “Pork barrel spending,” 16
 Positive correlation, 145–147
 Positive externalities, 59–61,
 103, 347, 362
 Positive incentives, 88
 Positively skewed distribution,
 142, 144
 Positive trend, 249
 Post-test-only control group
 experimental design,
 111–112, 112*f*, 113
 Powell, Colin, 37
 PPBS (Planning-Programming
 Budgeting System), 35, 62
 Predictions. *See* Forecasting;
 Projection techniques
 Prescription lag and delay in
 policy implementation, 67
 Presentation, effectiveness of,
 382–386
 Present value, 355–363
 choice of discount rate,
 360–362
 definition of, 356
 internal rate of return,
 362–363, 363*f*
 long-term vs. short-term
 solutions, 359–360
 President, U.S.
 agenda setting by, 77
 policy adoption
 considerations, 91–92
 Report of the President, 196
 Price deflator, 204
 Price increases in services,
 61–62
 Price indexes, 206
 Price makers, 53
 Price takers, 50, 53
 Primary data, 88, 177, 179–189.
 See also Sources of data
 Prisoner's dilemma, 337–338,
 338*t*
 strategies to overcome,
 339–341
 Private cost, 60
 Private goods, 6–7, 6*t*
 Privatization, 68, 69
 Probability, 149–176
 asymptotic, 151
 baseball, 152–153
 central limit theorem,
 155–156
 chi-squared test, 165–169
 comparison of different
 series of numbers,
 162–164
 confidence interval,
 159–161
 decision tree analysis and,
 323
 distribution, 151–157, 153*t*,
 154*f*
 hypothesis testing, 157–165
 independent events and, 150
 joint probability
 distribution, 167
 mutually exclusive events
 and, 150
 null hypothesis, 156–157
 number of tries and,
 150–151, 151*f*, 322

- objective probability, 149–169
- Pearson's r (correlation coefficient), 167
- risk vs. uncertainty, 320
- subjective probability, 169–170, 170*t*
- t distribution, 161–162, 284–285
- Probability-based sampling, 185–186
- Procedural justice, 121
- Production efficiency, 70
- Program termination, 96
- Project acceptance or rejection, 119–121
- Projection techniques, 225–318
 - adjustment methods, 256–260, 257*f*
 - analysis, elaborateness of, 378–382
 - analysis of historical data, 249–269
 - calculating the mean, 263
 - causal prediction vs., 225–226, 226*f*
 - choice of, 265
 - components of trend, 249–253, 250*f*, 253*f*, 256
 - Delphi technique, 232–237
 - expected utility model, 242–244, 243–244*f*
 - explaining the present with the past, 249–269, 311–313
 - feasibility assessment technique (FAT), 238–241. *See also* Feasibility assessment technique
 - fiscal impact analysis (FIA), 229–230, 229*t*
 - higher-order polynomials, 299–301, 300*f*
 - inadequate history available, 227–231
 - interval forecasts, 307–311, 310*f*
 - inverse forms, 302–303, 303*f*
 - irrelevant independent variables, 303–304
 - judgmental method of projection, 227, 232–245
 - least squares, 271–318
 - log-transformed forms, 301–302, 302*f*
 - moving average, 263–264, 264*f*
 - multiple regression, 272, 287–296
 - naïve projection of immediate past, 261–263, 262*t*
 - nonlinear trends, 296–305
 - patterns of time trend, 253–256
 - point forecasts, 306–307
 - polynomial forms, 296–299
 - seasonal adjustment, 256–259, 258*t*, 259*f*
 - simple regression, 271, 275–287, 277–278*t*. *See also* Simple regression model
 - single-factor projection, 227–229, 230–231, 231*f*
 - smoothing out fluctuations, 260–265
 - trend adjustment, 259–260, 260*t*, 261*f*
- Property values, 353–354
- Proportions test, 164–165
- ProQuest database, 195, 196
- Prospect theory, 14
- Public Affairs Information Service (PAIS)
 - PAIS International (online), 196
 - Public Affairs Information Service Bulletin*, 200
- Public goods
 - allocation of, 7–8, 53
 - care of, 9
 - costs of, 64–65
 - definition and classification, 6, 6*t*
 - ideological positions and, 18
 - Pareto principle applied to, 107
- Public policy curriculum, 236–237
- Public policy process, 75–100
 - agenda setting models, 76–87
 - behavioral models, 88
 - conflict, 89–91
 - costs and benefits analysis, 86–87
 - ethics and, 9–14, 40–42
 - forecasting, 88–89
 - goals of governmental interventions, 15–23, 15*f*
 - individualistic vs. collective societies and, 2
 - policy adoption, 91–92
 - policy change, 94–95
 - policy evaluation, 93–94
 - policy formulation, 87–91
 - policy implementation, 92–93
 - policy termination, 96–97
 - sensational events' impact on, 83–85
- Public services, charging for, 70–72
- Pure public goods, 6*t*, 7
- al-Qaida, 33
- Quade, E. S., 130
- Quadratic score, 169, 170*t*
- Quadratic trends, 253, 297, 298*t*, 299*f*
- Qualitative change, 288
- Quantification of survey data, 189–190
- Quantitative techniques, use of, 35–40
- Quasi-experimental research designs, 110, 112–115
- Questionnaires. *See* Survey instruments
- Quick decision making, 67–68, 144, 377–378

- R^2 (coefficient of determination)
 adjusted R^2 , 303–304
 calculation of, 280t
 defined, 279
 problems when high, 280
 $\bar{R}^2$, 303–304
Race Traits and Tendencies of the American Negro (Hoffman), 223
 Racial profiling, 212–213
 Rafool, Mandy, 128
 Raiffa, Howard, 344
 Randa, Laura E., 100
 Randers, Jorgen, 269
 Random errors
 in forecasting, 273–274, 274f
 in sampling, 184–185, 185f
 in trends, 249–250
 Random sampling, 185–186
 Range, 137–138, 139f, 159
 Ranking of resources, 240, 240t
 Raphael, David D., 73
 Rappaport, Anatol, 340
 “Rational Fools” (Sen), 5
 Rational fools, world of, 5–14
 Rationality, 1–26
 illogic of “rational beings,” 6–9
 and politics, 2–5
 un-rationality in political behavior, 9–14
 Rational model of decision making, 35
 Ratio scales, 178
 Rawls, John, 42, 108, 129
 Rawlsian criterion, 108, 119
 Reagan, Ronald
 defense budget and, 221
 federal assistance to state and local governments, 292
 federal regulations and, 11
 job performance of, 12
 navy ships, 64
 privatization and, 69
 rational decision making and, 35
 on social welfare inefficiency, 17
 Real income, 206
 Real terms, 207
 Reason, rationality, and public policy, 1–26
 Recognition gap and delay
 in policy implementation, 67
 Rector, Robert, 40, 41
 Reggio, Michael H., 100
 Regression analysis, 301
 autoregressive model, 311
 causal regression model, 353–354
 multiple regression, 272, 287–296. *See also* Multiple regression
 simple regression, 271, 275–287, 277–278t. *See also* Simple regression model
 Regression equation, 275
 Regression toward the mean, 124
 Regulatory action for public good, 9
 Reiter, S., 280, 317
 Relative frequency, 133, 149
 Relative measure of efficiency (opportunity cost), 16
 Rent seeking, 18, 56–57, 57f, 118–119
Report of the President, 196
 Representativeness, 122–125, 181
 Required compensation method, 352
 Research design. *See* Critical thinking and research design
 Research institutes (think tanks), 28
 Retrievability, 125–126
 Revealed preference, 184
 Reverse U-shaped pattern, 253–254, 254f
 Ridley, Jasper, 44, 100
 Risk
 required compensation method in risky jobs, 352–353
 tolerance, 330–332, 331f
 uncertainty vs., 320
 Risk averters, 104, 331
 Risk neutral, 331
 Risk takers, 104, 331
 Rivalrous quality (exhaustibility) of a good, 6, 6f
 Rivlin, Alice, 39, 45
Roe v. Wade (1973), 68–69
 Roelithsberger, F.J., 183
 Roosevelt, Franklin D., 10, 38, 65, 68, 181
 Ross, Jennie-Keith, 98
 Ross, Marc H., 98
 Rousseau, Jean-Jacques, 105
 Row marginal, 166
 Rubin, P. H., 99
 Rubinfeld, Daniel L., 176
 Rumsfeld, Donald, 33, 37
 Sabatier, Paul, 92, 98, 99, 100
 Sabine, George, 105, 129
 Sackman, Harold, 248
 Sageman, Marc, 116, 129
 St. Claire, G. K., 99
 Saint-Germain, Michelle A., 248
 Samet, Jonathan N., 371
 Sample error, 187
 Sample population, 180, 185–186
 Samples, 133
 Sample size, 123
 Sample standard deviation, 140
 Sampling theory, 180, 185–186
 San Diego, California
 garbage collection fees, 72
 illegal immigration, 231
 racial profiling in police stops, 212

- sports stadium, public
 - financing of, 104
- San Diego Union Tribune*, 212
- Sapolsky, Harvey, 130
- Satisficing behavior in agenda setting, 83
- Sawhill, Isabel, 41
- Sawicki, David S., 129
- Scatter plots, 210
- Schattschneider, E. E., 99
- Schick, Allen, 130
- Schmid, Alex P., 84, 99
- Schools
 - benefits of free public education, 70
 - California immigration controversy in, 51
 - Drug Abuse Resistance Education (D.A.R.E.), 88
 - higher education costs, 52, 60–61
 - safety and violence, 84
 - testing, programs to improve performance, 93–94
 - vouchers, 113
- Schumpeter, Joseph, 25
- Schwartz, Eli, 344
- Scientifically based public policy, 373
- Search engines, 194
- Search goods, 57, 59
- Searching for information, 126, 193–200, 194f
- Seasonal adjustment, 256–259, 258t, 259f
- Seasonal trend, 251, 252f
- Seat belt use on Texas school buses, 367–368
- Secondary data, 192–200
 - defined, 177
 - hypothesis testing, use in, 88
 - online searches, 193–200, 194f
 - searching for information, 193
- Securities and Exchange Commission (SEC), 11, 11f
- Security as policy goal, 15, 15f, 21–22
- Security-liberty trade-off, 22–23
- Self-interest
 - Congress, when voting on legislation, 87
 - as determinant in agenda setting, 87
 - as human motivation, 5–14
- Self-selection bias within the sample, 181–182
- Self-utility-maximizing economic being (*homo economicus*), 8, 9, 67
- Self-utility-maximizing human being, 6
- Sen, Amartya K., 5, 23, 24, 26, 99, 343
- Sensational events' impact on agenda setting, 83–85
- Sensitivity analysis, 364
- Service costs, 61–62
- Service delivery systems, 180
- Shadow price, 360
- Sherman Act of 1890, 54
- Shiller, Robert, 67, 68, 75
- Significance tests
 - hypothesis testing and level of significance, 157
 - for individual coefficients, 283–286, 286t
- Silverstein, Arthur M., 343
- Simon, Dorothea, 248
- Simon, Herbert, 98, 248, 376
- Simple regression model, 275–287
 - accuracy of results, 278–280
 - estimation results, 286
 - high R^2 , 280
 - linear trends, use in, 271
 - number of observations needed for, 286–287, 287f
 - population over time, 277t
 - relevance of estimated coefficients, 280–283
- significance of individual coefficients, 283–286, 286t
- straight-line regression equation calculations, 278t
- Simplicity and pluralist model of agenda setting, 80
- Simulation of the market, 68, 69–72
- Singer, David, 198
- Singh, Harinder, 176, 387
- Single-factor projection, 227–229, 230–231, 231f
- Size of sample, 186–187
- Skewed distribution, 142
- Skewness and symmetry of distribution, 141–144
- Slope dummy, 294, 294t
- Small, Melvin, 198
- SMCRA (Surface Mining Control and Reclamation Act), 91
- Smirnov, Oleg, 26
- Smith, Adam
 - invisible hand, 21
 - mercantilism, 47–48
 - rationality, 3, 5, 24
 - self-correcting market, 52
- Smoking
 - social costs of, 60
 - teen smoking, 79–80
- Snail darter, 121
- Social costs and benefits, 59–61, 346–347, 350, 350f, 351
- Social objectives, 106
- Social Security, 17, 267
- Social significance and pluralist model of agenda setting, 80
- Social time preference, 361
- Social tyranny, 21
- Social welfare
 - designing research for, 118
 - problem of defining, 39, 63, 107–108, 117

- Solomon four-group experimental design, 110–111, 111f
- Sources of data, 177–202
- accuracy of measuring scales, 178–179, 179f
 - choosing the sample population, 185–186
 - focus groups, 190–192
 - online searches, 193–200, 194f
 - polling methods, 189
 - primary data, 88, 177, 179–189
 - print publications, 200
 - quantification of survey data, 189–190
 - random errors in sampling, 184–185, 185f
 - reporting of survey results, 190
 - secondary data, 88, 177, 192–200
 - size of sample, 186–187
 - survey instruments, 187–189
 - systematic errors in sampling, 180–184, 181f
 - terrorism research, 198–199
 - types of measurement, 177–178
 - unavailable data, 200
- Soviet Union, 245
- Space program, 292, 361
- Specific costs and benefits, 86–87, 86t
- Specificity and pluralist model of agenda setting, 79–80
- Split policy change, 95
- Sports stadiums, financing of, 101, 103–106, 105t, 117–118
- Spotted owl, 353
- Squared deviation, 275
- S-shaped trend, 254–255, 255f
- Stabilization, 53
- Standard deviation, 140
- population standard deviation, 187, 281
 - of probability distribution, 152
 - sample standard deviation, 140, 187
- Standard normal distribution, 152, 154, 155f
- Starr, Paul, 197, 202
- Stationary series, 251
- Statistical Abstract of the United States*, 199
- Statistically significant coefficients, 284–285
- Statistics, 131–148
- behavioral model validated by, 88
 - central limit theorem, 155–156
 - central tendency, 131, 132–137, 144
 - deceptive use of, 213–222
 - dispersion, 131, 137–144, 138f, 138t
 - government publications of, 197
 - grouped data, 136–137, 137t
 - interpretation of, 214–221
 - lagging, 261, 263
 - methods of descriptive statistics, 131–147
 - persuasion by, 384–385
 - pictorial presentation of, 218–219f, 218–220, 220t
 - presentation of, 203–224
 - printed publication of government data, 200
 - skewness and symmetry of distribution, 141–144
 - stem-leaf method, 144–145, 145f, 145t
 - tabular presentation of, 221–222, 222t
 - techniques used for analysis, 42–43
 - weighted estimates, 135–136, 136t
- STAT-USA, 199
- Stem-leaf method, 144–145, 145f, 145t
- Stewart, Joseph, 98, 100
- Stiglitz, Joseph, 99
- Stimulation of the market, 68–69
- Stockholm International Peace Research Institute, 21
- Stokey, Edith, 117, 130, 363, 371
- Stone, Deborah A., 26, 99
- Stonecash, Jeffrey, 148
- Stratified sampling, 186
- Strikes, 334–335, 337
- Structured conflict, 234
- Student's *T* distribution, 161–162, 284–285
- Subgovernmental model of agenda setting, 82, 82f
- Subjective judgment, 39–40
- Subjective probability, 169–170, 170t
- Summing the errors, 275, 275f, 279
- Surface Mining Control and Reclamation Act (SMCRA), 91
- Survey instruments, 184, 187–189
- Surveys, 180
- errors in. *See* Random errors; Systematic errors in sampling
 - quantification of data, 189–190
 - reporting of results, 190
- Svara, James H., 44
- Swine flu, 319–320, 322, 323–324f
- Symbionese Liberation Army, 9
- Symmetric distribution
- asymmetric distributions compared, 143t
 - asymmetric series compared, 143f
 - data showing, 142t
 - distribution with extreme value, 135f
 - example of, 142f

- mean, median, and mode, 144t
- shapes of, 152
- skewness and, 141–144
- statistics, 134, 134f
- Systematic errors in sampling, 180–184
- bias in survey instrument, 184
- collection bias, 182
- example of, 181f
- lying by survey respondents, 182
- nonrepresentative sample, 181
- observation errors, 182–184
- self-selection bias, 181–182
- Systemic (noninstitutional) agendas, 76–77
- Tajfel, Henri, 12–13
- Target outcome, 150
- Taxation
- California Proposition 13 (tax revolt), 70, 332–333
- increased revenues as benefit from subsidized sports stadium, 103
- Taylor, Charles, 182, 202
- T* distribution, 161–162, 284–285
- Teen smoking, 79–80
- Telephone interviews, 189
- Temporal relevance and pluralist model of agenda setting, 80
- Tennessee Valley Authority (TVA), 121
- Termination of policy, 96–97
- Terrorist attacks
- 9/11 terrorist attacks, 21, 22, 33
- on U.S. embassies in East Africa, 33
- on U.S.S. *Cole*, 33
- Texas school buses and seat belt use, 367–368
- Thayer, Frederick, 44
- Thornton, Robert J., 344
- Time
- effect on research design, 114–115, 273, 273f
- governmental policy, timing of, 15–16, 67
- patterns of time trend, 253–256
- present value analysis, 355–363
- Time*, 20
- Time horizon, 359–360, 360f
- Time preference, 356, 357f
- Time-series data
- graphical presentation of, 211
- method of least squares for trend projection, 271–272
- Tit-for-tat strategy, 340
- Tobacco Institute, 28
- Toll goods, 6t, 7
- Torgerson, Douglas, 45
- Trade. *See* Exports and imports
- Trade-offs in public policy, 15–16, 15f, 19, 22–23
- Trade secrets, 57
- Traffic, 75, 94, 119–120, 212, 353–354
- Tragedy of the commons, 9
- Transportation and highway construction, 366
- Trend adjustment, 259–260, 260t, 261f
- Trend coefficient, 276
- Trends, 225
- abrupt changes in, 288–292, 288t, 289f, 290f
- analysis, 115, 249–253, 250f, 253f, 256
- multiple regression analysis, 272
- simple regression analysis, 271, 275–287, 277–278t. *See also* Simple regression model
- Truman, David B., 98
- Trust, 338–339, 341
- TVA (Tennessee Valley Authority), 121
- Tversky, Amos, 14, 26, 122, 124, 125, 130
- Two-tailed test, 160–161, 161f, 162
- Type I error, 167
- Type II error, 167
- Tyranny of the majority, 21
- Uncertainty, 38, 320–321, 338–339
- Unconditional forecasting, 306–307
- Undocumented aliens. *See* Illegal immigration
- Unemployment, 55–56
- Unintended consequences, 109, 110f
- United Nations
- Children's Fund, 28
- Commission on Human Security report of 2003, 23
- Development Programme report on human security of 1994, 23
- statistical publications, 197
- United States Institute of Peace, 28
- Universal Declaration of Human Rights, 37
- "Un-rational being" and political behavior, 9–14
- Upper hinge, 137
- Ury, William, 341, 344
- U.S. Army Corps of Engineers (COE), 325, 343
- USA PATRIOT Act of 2001, 32
- U-shaped pattern, 253–254, 254f
- Utilitarians, 106
- Valuation of costs and benefits, 349t, 350–355
- Value conflict, responses to, 32–34, 34f, 36–37

- Variables, 133–134, 294, 294*t*
 See also Dependent variables; Independent variables
- Variable weight, 205
- Variance, 139–140, 140*t*
- Vaupel, James, 377, 387
- Vertical shifts, 288
- Vining, Aidan, 32, 34*f*, 43, 44, 45
- Violence
 cost-benefit analysis for gun control, 86
 political division over gun control, 89
 school violence as example of pluralist model of agenda setting, 84
- Voting and limits of democracy, 63–64, 64*t*, 65*t*
- Wall Street Journal*, 221
- Walmsley, Julian, 344
- War on terror, 33
- War Powers Resolution of 1973, 66
- Washington Post*, 36
- Watkins, Sherron, 36–37
- The Wealth of Nations [An Inquiry into the Nature and Causes of the Wealth of Nations]* (Smith), 3, 47, 49
- Wealth transfer, 104
- Weather Underground Organization, 9
- Web searches, 193–200, 194*f*
- Weighted estimates, 135–136, 136*t*
- Weimer, David, 32, 34*f*, 43, 44, 45
- Weinberger, Caspar, 221
- "Welfare queen," 17
- Westen, Drew, 12, 14, 25
- Wheelwright, Steven, 269
- Whistleblowers, 34, 36–37
- Wildavsky, Aaron, 35–36, 45
- Will, Susan, 344
- Williamson, Joel, 224
- Wilson, Woodrow, 30, 44
- Woods, B., 99
- Woodward, Bob, 36
- World Bank statistical publications, 28, 197
- World Factbook*, 197, 309
- World Handbook of Political and Social Indicators*, 182
- World Health Organization, 78
- World Meteorological Organization, 28
- World War II battle and use of game theory, 335–337, 336*t*
- Wright, George, 176
- X-inefficiency, 54
- $\hat{Y}$ ("Y hat one"), 274
- Yates, J. Frank, 176
- Yearbook of International Trade Statistics*, 197
- Yergin, Daniel, 73
- Yoo, John, 37
- Zagare, Frank C., 344
- Z distribution, 158
- Zeckhauser, Richard, 117, 130, 363, 371
- Zero-sum games, 338–339
- Ziegler, Harmon, 81, 98
- Zone of ignorance, 107
- Z score, 158, 159, 162, 164, 186
- Z test, 158–159
- Zupan, Mark, 91, 99

Images have been losslessly embedded. Information about the original file can be found in PDF attachments. Some stats (more in the PDF attachments):

```
{
  "filename": "NDA3NDgzNTEuemplw",
  "filename_decoded": "40748351.zip",
  "filesize": 253910892,
  "md5": "a938ffd4d84e693573637a6a4ff33c3e",
  "header_md5": "e4c14a73c53e4126800ce9af4b6f7b57",
  "sha1": "384b4b410b489f9d4853606fd2a4e0d37a83d6e0",
  "sha256": "6479ad744d1f487f6e07084fe38b32262df807dcb7edf625b3728a7c5081a64a",
  "crc32": 2749717635,
  "zip_password": "",
  "uncompressed_size": 270107195,
  "pdg_dir_name": "ANALYZING PUBLIC POLICY CONCEPTS,TOOLS,AND TECHNIQUES SECOND EDITION_40748351",
  "pdg_main_pages_found": 470,
  "pdg_main_pages_max": 470,
  "total_pages": 493,
  "total_pixels": 2741614412,
  "pdf_generation_missing_pages": false
}
```